DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A DYNAMIC THEORY OF REGULATION

Burton H. Klein

SOCIAL SCIENCE WORKING PAPER 199

February 1978

CALIFORNIA INSTITUTE OF TECHNOLOGY

Division of the Humanities and Social Sciences
Pasadena, California 91125

A DYNAMIC THEORY OF REGULATION

Burton H. Klein

Social Science Working Paper
Number 199
February 1978

A DYNAMIC THEORY OF REGULATION

Burton H. Klein
California Institute of Technology

To date, discussions of economic regulation have concentrated on assessing the impact of regulation on static efficiency. Fascination with statical analyses has stemmed from two important factors. First, on the basis of mathematics developed long ago for the physical sciences, regulatory arguments can be stated in such highly rigorous terms that economics assumes the appearance of a successful science. Second, economists continue to regard the state of Pareto optimality as a special kind of heaven in which all truths are known, and all problems are completely solvable. From this perspective a regulatory bias, such as the Averch-Johnson effect,[1] which causes investment decisions to be more capital intensive, is looked upon as a sin comparable to that committed by the infidels, and a holy crusade is mounted to ascertain what can be done to restore the economy to its heavenly virtues.

More recently, in "The Theory of Economic Regulation,"[2] George Stigler has made a more serious attack on economic regulation. By pointing to the effect of regulation on limiting entry, he has singled out a more important and observable effect of regulation than the Averch-Johnson effect. And by arguing that this effect results from the ability of a rather small number of firms to redistribute income in their own favor, he has pointed to a regulatory consequence that cannot be eliminated by tinkering with the regulatory mechanism.

By contrast, the dynamic theory of regulation presented in this paper is to be regarded as an earthly, rather than a heavenly, theory. Like Stigler, I refuse to believe that economic regulation typically occurs because of a market failure. But, unlike him, I do not take the concept of a competitive equilibrium (involving, among other things, a world of perfect knowledge and completely unimaginative entrepreneurs) as a norm: not because such a norm is seldom approximated in the real world, but, rather, if their environment is sufficiently challenging, a few firms managed by imaginative entrepreneurs can be counted upon to do better than an infinite number of firms managed by computers. Unfortunately, however, firms do not always cooperate with each other by engaging in rivalry to generate cheaper or better alternatives. In fact, the older and less adventurous firms in an industry can develop a common interest in protecting themselves from such rivalry.

According to the theory presented in this paper, economic regulation does not ordinarily result from a market failure; nor does it ordinarily result from an attempt to divide monopoly profits among a smaller, rather than a larger, number of firms. It results from a difference in utility functions with respect to both individuals and firms: for example, as when older firms in an industry lose their taste for rivalry and are able to obtain government assistance to impose "order" on the industry.

Another difference between the theory presented here and Stigler's is that I do not believe it necessary to assume that regulatory agencies are captured by regulated industries. According to my theory, regulatory agencies and regulated industries make good partners because both desperately want an environment with a minimal degree of unpredictability.

Finally, dynamic theory measures the impact of regulation in a different way from static theory. Dynamic theory is mainly concerned with the impact of regulation upon dynamic efficiency, that is, on the rate firms are likely to generate better or cheaper alternatives. And, according to dynamic theory, regulation has the effect of making an environment more certain at the expense of making future reductions in costs and prices more uncertain. Therefore, if my argument is correct, in a relatively few years the dynamic costs of regulation will exceed the static costs.

Dynamic behavior, as defined in modern (i.e., nondeterministic) terms, differs from static behavior in two important respects.[3] First, whereas static theory assumes that behavior can be predicted on the basis of initial conditions, dynamic theory assumes that on the basis of feedback entrepreneurs can use their imaginations to make quite unpredictable changes in initial conditions. Second, and related, whereas static processes result in reversible change, dynamic processes result in irreversible change -- that is, they change the probability distributions of the world. Changes resulting from movements in factor prices, for example, are quite reversible. On the other hand, the numerous discoveries made during the process of pioneering the automobile -- discoveries which permitted almost an order of magnitude reduction in costs between 1900 and 1925 -- were not only unpredictable when viewed as isolated events, but were irreversible processes which changed subjective probability distributions of the world.

Dynamic behavior is inescapable because nature places limits on what can be achieved within existing technologies. For example, consider Sadi Carnot's theory about the efficiency of machines.[4] It tells us that the physical efficiency of a machine -- the amount of work an engine can generate in relation to the heat put into it -- is proportional to the temperature difference which can be obtained, or:

$$\text{Physical efficiency} = \frac{W}{Q_1} = \frac{T_1 - T_2}{T_1}$$

where $\frac{W}{Q_1}$ is the output of work per unit of heat, and $\frac{T_1 - T_2}{T_1}$ is obtained by taking the difference between the highest and lowest temperatures at which the engine runs and dividing that difference by the highest temperature.

What limits the temperature difference is (1) the best available fuels, (2) the best available materials, and (3) the engineering know-how to make the machine approximate an ideal reversible machine. Overcoming these limits makes significant improvements in the efficiency of a machine highly unpredictable. To be sure, the engineer knows which way is up; that is, he knows, if successful, certain measures will increase the efficiency of a machine. But, before the crucial experiments are made he has no way of predicting how much the physical efficiency can be increased.

This is not to say, of course, that major improvements in efficiency come about only through discovering new technologies. Organizational changes are important if new technologies are to be effectively utilized.

People accustomed to thinking in terms of equilibrium economics tend to regard static behavior as normal and dynamic behavior as abnormal. However, the ability to engage in dynamic behavior constitutes the single most important difference between the lower and the higher primates. For example, when confronted by necessity, chimpanzees have displayed an amazing ability to discover new tools for themselves. And there is no reason to suppose that real-world entrepreneurs are any less imaginative than chimpanzees.

The first section of this article will present three general propositions concerning the relationship between the degree of uncertainty in an environment and the propensity for engaging in dynamic behavior in order to reduce risk and uncertainty. The second will present a conceptual model for more clearly understanding how, by making an economic environment more certain, regulation reduces the incentive for dynamic behavior and seriously constrains the entrepreneur's freedom of choice for dealing with uncertainty. And in the third section I will argue that in a relatively few years the dynamic impact of regulation will likely dwarf the static impact.

I. Uncertainty and Risk-taking

The principal differences between a world of static efficiency and a world of dynamic efficiency are to be found in the nature of the uncertainties involved and the manner with which entrepreneurs deal with them. Let us consider three worlds: a world of zero uncertainty, a world of weak or statistical uncertainties, and a world of strong uncertainties. A good example of the first is the completely deterministic world described by general equilibrium economics: a world that is more deterministic than the physical world. According to the uncertainty principle in physics, making an observation with respect to the position and momentum of a particle affects the phenomenon in question.[5] But, according to Walrasian general equilibrium economics, the entrepreneur is assumed to know the true trading price before any trading takes place.[6] A world of weak uncertainties is one which can be described entirely in terms of probability distributions. True, in such a world there still might be substantial differences in subjective probability distributions. And such differences plus large transaction costs may result in a world wherein risks are not completely insurable. But as a polar case, let us assume a world with small differences in subjective probability distributions and negligible transaction costs: a world in which insurance companies are formed to provide insurance against every imaginable risk. If a house burns down while the owner is at work an indistinguishable replacement is guaranteed by the time he returns home. If his wife should run away with another man the insurance company guarantees to provide him with a quite satisfactory substitute. And firms can obtain 100 percent insurance against losses of markets to their competitors.

A world of strong uncertainties is one in which it is not possible to obtain 100 percent insurance against all conceivable risks. How do individuals and firms deal with a world of strong uncertainties? Generally speaking, negative feedback plays the role of stimulating a search for new alternatives in order to reduce the risk in question. For example, negative feedback in the form of a declining market share will indicate that a firm has not been as successful as its rivals. And the first step involved in generating a more satisfactory alternative consists of deeply understanding why the current alternative was unsatisfactory.

What can be said about the incentives associated with worlds of zero, weak, and strong uncertainties?

<u>Proposition I</u>: From the point of view of the incentives provided and the behavior involved, there is no fundamental difference between a world of zero uncertainty and one of weak uncertainties. The very act of forming insurance companies is to convert strong uncertainties for individuals into weak uncertainties for a society of insured individuals. If we assume that all hazards are completely insurable, then with 100 percent insurance people would be indifferent as to whether the events insured against actually occurred. Thus, in neither world is genuine risk encountered nor is there an incentive to heed feedback -- or to display the imagination of an enterprising chimpanzee.

Proposition II: It is assumed that, generally speaking, business firms are risk averse. But it is also assumed that firms must always compare two risks: the technological risk involved in developing a new product with the risk of a loss in the share of the market to a rival. It is further assumed that the greater the market risk associated with the introduction of new products, the larger the technological risks firms must be willing to take if they hope to survive. In other words, it is assumed that not only is a hidden hand in operation, but also a hidden foot, whose threat depends on the degree of rivalry involved. In the classical environment of weak or zero uncertainties the hidden foot plays no role whatsoever, but as the uncertainties become stronger the incentive to become involved in risk-taking becomes greater. It is also assumed that business firms can employ more or less imaginative people, and that the degree of imaginativeness will depend upon the degree of uncertainty associated with the environment in question.

On the basis of this argument, my second Proposition can be succinctly stated as follows: the smaller the degree of uncertainty associated with an environment (or some particular aspect of it), the closer it will resemble the case of 100 percent insurance against risk, and, consequently, the poorer will be the incentives for dynamic behavior. Conversely, the stronger the uncertainties associated with an environment, the greater will be the incentives for dynamic behavior. This proposition is similar to the "moral hazard" observation made by Arrow in his famous insurance essay.[7] However, my argument differs

from Arrow's in two important respects. In the first place, whereas Arrow was concerned with only the relationship between insurance companies and the insured, it is my conviction that the logic of the argument holds in all cases in which collective actions of individuals or firms make an environment seem more certain. If greater insurance coverage against fires weakens the incentive to take risks to prevent a fire from occurring (the moral hazard point), then, why should a cartel aimed at preserving each firm's share of the market not have the same impact on the willingness to take risks in order to introduce better or cheaper alternatives? Is not a cartel an arrangement for insuring firms against competitive risks? Conversely, when firms cannot so insure themselves against the actions of the hidden foot, and they want to make their survival more predictable, must they not be prepared to take risks?

Secondly, Arrow does not explicitly deal with changes in behavior when incentives change. Implicit in Proposition II is the assumption of a continuum in the degrees of uncertainty and in the dynamic behavior required to deal with uncertainty. Relatively few, if any, of man's activities involve dealing with absolutely no uncertainty. For example, even in such a routine activity as pouring a cup of coffee feedback is taken into account by tilting the hand holding the percolator to avoid overfilling the cup. But in this case the uncertainties are weak; and, in my terms, the behavior involved would be described as cybernetic rather than dynamic. Likewise, a sea captain who makes course corrections by observing the interactions between the waves and his ship is engaged in cybernetic, rather than in dynamic, behavior. To engage in dynamic behavior the captain would have to react to conditions never before encountered.

Next, imagine a world whose probability distributions are changing ever so slowly: a Bayesian world in which, as a result of new experiments, probability distributions change more or less continuously. This might seem like a world in which no imagination is required. But is it? Ordinarily when an entrepreneur buys information to reduce variances he is borrowing an idea used earlier in another application. But as much as a borrower of ideas would like to look before he leaps, it is practically impossible to borrow ideas successfully without running into unexpected problems, unlearning something the entrepreneurs previously believed was true, and without having to generate a new hypothesis.[8] In other words, like it or not, borrowers do generate irreversible change; and, consequently, the distinction between "borrowing an idea" and "generating an idea" is a matter of degree and not of kind.[9]

How does an entrepreneur generate an hypothesis previously assigned a probability of zero? Imagination is required to disassociate hints from particular experiments or experiences so they can be used to suggest new ideas. In fact, it is this ability of the higher primates to utilize hints to make new discoveries that enables them to adapt to new circumstances. These hints may come from a narrow or a broad group of experiences. And, in general, the larger the advances to be made, the more randomness that is required in the hints.

For example, in 1904, Henry Ford knew that the name of the game, if he was to capture first place in sales from Buick, was to produce a car which was both durable and inexpensive. He was also aware that achievement of that goal would require hundreds of inventions -- from the creation of stronger materials to discovery of methods for obtaining more efficient production lines. How did he go about attaining these innovations? To obtain a great deal of randomness in hints Ford hired people from a wide diversity of backgrounds: from a German metallurgist who assisted by producing stronger steel, to a manual arts teacher who thought of the idea for moving production lines. By 1925, of course, the name of the game had shifted from static to dynamic efficiency; but by then, Ford, as well as other automobile companies, were optimized for a lower degree of uncertainty: there was less diversity in the people hired; and as the organization became more structured and more statistically efficient there was less randomness in their communications.

In short, there is an entire range of possibilities for adjusting organizations to deal with higher or lower degrees of uncertainty, and for engaging in higher or lower degrees of dynamic behavior.

What factors determine the degree of uncertainty in an industry? As was indicated above, one factor is the degree of rivalry between firms: the larger the advances firms aim to make, the greater the degree of uncertainty. What determines the degree of rivalry deserves a good deal of further study. However, it is my conviction that one major factor is the entry of new firms into the industry in question. As long as new firms continue to enter an industry, progress is likely to be rapid; when entry becomes closed, progress is likely to slow down. For example, in the automobile industry many of the important advances were made by relative newcomers to the industry: the leading firms in 1900 were by no means the leading firms in 1910; and the leading firms in 1910 were by no means the leading firms in 1925.[10] In the fields of aircraft and aircraft engines much the same was true. For example, before the advent of the DC-2 Douglas had never developed a commercial airplane. The firms that pioneered air-cooled aircraft engines after World War I were not those that pioneered liquid-cooled engines; and those that later pioneered the jet engine were not involved in pioneering reciprocating engines. And, according to John Tilton, more or less the same thing happened in the more recent development of semiconductor technology.[11]

Suppose that technology in a particular industry has been relatively stagnant for some years, and a discovery is made which has the effect of inaugurating a period of rapid progress, for example, the oxygen process in steel, the jet engine, or synthetic fibers. What kind of firms are responsible for making the industries' revitalizing discoveries? Of some fifty cases I looked into, not one discovery came from a major firm in the industry after there had been a significant slowing down in the rate of progress in that industry.[12]

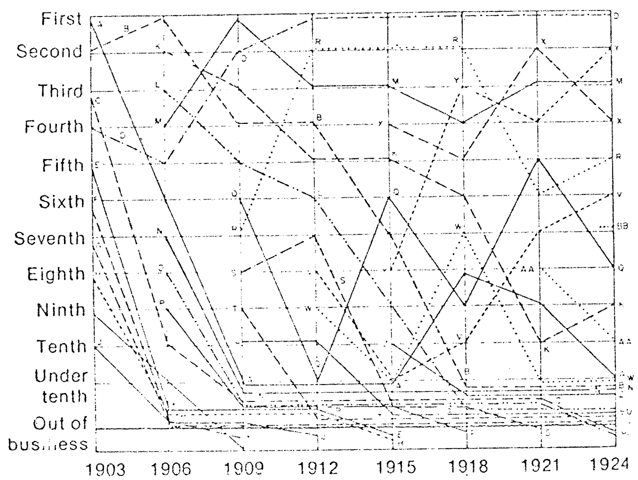
In short, entry of new firms plays a quite different role in dynamic theory from that played in static theory. According to static theory, firms make identical products at identical costs. The function of the entry of new firms, therefore, is to eliminate

monopoly profits and correct a misallocation of resources. On the other hand, the very essence of dynamic competition is the generation of better or cheaper alternatives. And the function of entry is to maintain an environment which, by featuring a high degree of risk and uncertainty, favors the generation of a wide diversity of ideas.

Generally speaking, when progress is rapid we can expect to observe not only large changes in market shares, but also quite unpredictable changes -- unpredictable inasmuch as they involve the entry of newer firms as well as the exit of older firms. For example, there is no question that progress in improving automobile technology was more rapid during the period 1900 to 1925 than during the period 1925 to 1940. During the former period, the average price of cars, as measured in real terms, was reduced about 75 percent -- while at the same time there was a substantial increase in quality. On the other hand, during the period 1927 to 1941, there was a very definite slowdown in the rate of progress -- measured either by reductions in cost or improvements in quality. As Figure 1 shows, this slowdown was accompanied by a substantial decline in feedback -- measured in terms of changes in market shares. To be sure, changes in the market shares of the ten leading automobiles are easier to predict today than they were during the period 1925 to 1940. But, relatively speaking, the period 1900 to 1925 was one of utter confusion. After being the leader in 1903, automobile A never again regained first rank; B and G declined from second place to below tenth place; and automobiles C, E, G, I, and J, which also ranked among the top ten in 1903, were made by firms that shortly

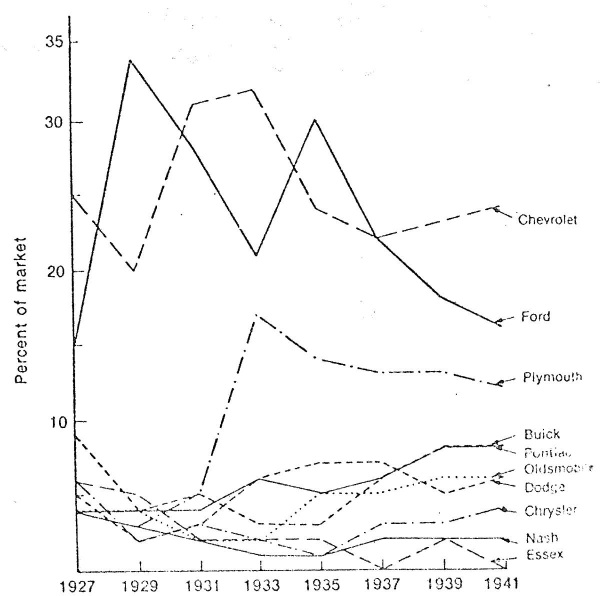
FIGURE I

FEEDBACK IN THE AUTOMOBILE INDUSTRY



Changes in Market Shares 1903 - 1924

Source: Burton Klein, Dynamic Economics, p. 100



Changes in Market Shares 1927 - 1941

Source: Ibid., p. 108

went out of business. In fact, of the ten leading automobiles in 1924, only three were produced by firms in business prior to 1924. Thus, it is apparent that when rapid progress was being made in the automobile industry changes in market shares were larger and more unpredictable than after slow history set in.

Proposition III: When incentives favor a high degree of risk-taking the greater will be the emphasis on dynamic efficiency; when they favor a low degree of risk-taking the greater will be the emphasis on static efficiency. Of course, every firm would like to make as much money as possible, but the way each achieves this goal will differ depending upon the relative emphasis placed on static and dynamic efficiency. In the pure case of static efficiency the enterpreneur takes his alternatives as given, and maximizes his profits subject to a predetermined menu of alternatives. On the other hand, in the case of dynamic efficiency the entrepreneur never takes his initial alternatives as given. He assumes a world of imperfect knowledge in which it is possible to discover better or cheaper alternatives. And, depending on the emphasis placed on either static or dynamic efficiency, the changes in the menu of alternatives can be marginal or highly discontinuous.

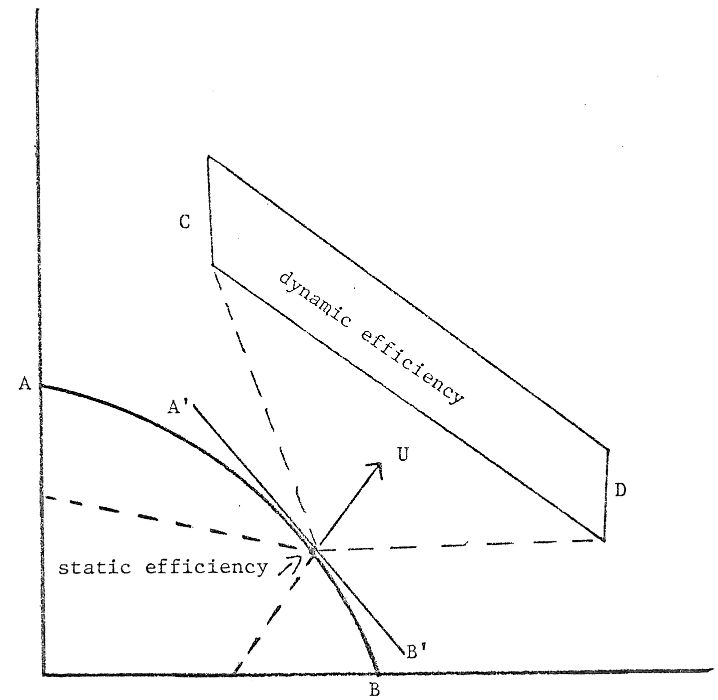
Implicit in Proposition III is the assumption that entrepreneurs cannot simultaneously enjoy a high degree of static and dynamic efficiency. For example, suppose that AT&T had been composed of only a Bell Telephone Laboratory and not a Western Electric. With the wide diversity of personnel and informal organization that characterized BTL twenty-five years ago, many marvelous inventions would have occurred -- but at the expense of static efficiency, that is, at the expense of a reliable telephone system. Conversely, with as tightly structured an organization as Western Electric has always been, a high degree of static efficiency would have been attained at the expense of advances in telephone technology occurring at a snail's pace.

A graphical representation of the trade-off between static and dynamic efficiency is provided by Figure 2. The curve AB is a technological frontier. It shows how, on the basis of the best available technology, two technological characteristics can be combined, say, the weight of an aircraft engine in relation to its thrust, and fuel consumption in relation to its thrust. The relative price (A'B') an entrepreneur will pay for these characteristics in order to maximize efficiency will, of course, depend upon their utility in given applications. For example, to develop a jumbo jet the entrepreneur would be willing to pay a relatively high price in terms of engine weight in order to minimize fuel consumption. Carefully selecting the point of maximum efficiency on a trade-off curve is the essence of static efficiency.

On the other hand, if the goal were dynamic efficiency the entrepreneur would be concerned with extending the technological frontier to the region CD. To be sure, he might like to have his cake and eat it too, that is, to simultaneously achieve the advantages of specialization and a large advance in technology. However, demanding a high degree of specialization is equivalent to imposing constraints on his ability to deal with uncertainty. And, assuming that the entrepreneur's ability to deal with uncertainty is finite,

FIGURE II

STATIC VERSUS DYNAMIC EFFICIENCY



imposition of such a constraint would make the cost of pushing out the frontier more expensive than need be. So, to be dynamically efficient the entrepreneur cannot demand a high degree of specialization. Conversely, the entreprenuer cannot demand a high degree of static efficiency without destroying his ability to deal with uncertainty.

II. A Conceptual Model of Regulation

Much work remains to be done before the following theory can be put forth in sophisticated mathematical terms. However, the first task in devising a predictive dynamic theory is to visualize, best one can, the changes in incentives and behavior which accompany regulation. To be sure, one can proclaim, as many economists have, that regulation occurs because of the need to deal with externalities. Or, like Stigler, one can proclaim that regulation occurs because of the ability of the dominant firms in an industry to secure the assistance of the government in limiting entry. But, no matter how much or how little mathematical sophistication is employed in devising such theories, they are to be regarded as "debating" rather than "scientific" theories. As scientists understand the meaning of science, the ultimate test of a theory is its ability to predict. Indeed, it can be predicted that in the absence of such a test scientists would still be debating whether the earth were round or flat (and, in the process, developing evermore devilishly complicated mathematics).

Why Are Economic Regulations Adopted?

Traffic lights represent a form of regulation. Why did cities begin to install them early in this century? There is no

doubt that traffic signals can allow traffic to flow more smoothly, thereby permitting an economy of scale in the utilization of streets that could not otherwise be realized. However, this argument can hardly explain why traffic signals were adopted while relatively few cars were being used. The more likely reason is that not all drove in the same manner: if all drivers had been highly cautious types, there would have been no need for traffic signals; on the other hand, if all had been highly adventurous drivers no traffic signals would have been wanted. To be sure, the interests of pedestrians also had to be considered. But, again, we must observe that the interests of pedestrians and drivers differed: while farmers living in North Dakota probably had no keen interest in organizing a crusade for traffic signals, timid pedestrians in Boston assuredly thought differently.

It is my prediction that more or less the same is true of economic regulation: regulation occurs because of the dominance of one group of interests over another less powerful group. When an industry is in the process of becoming established not all firms have the same ability to deal with uncertainty. To exploit new discoveries firms must become better optimized for static efficiency; but, by becoming more structured inevitably some of their ability to deal with uncertainty is lost. Hence, firms entering the industry, say, ten years later, can have a decided advantage in dynamic efficiency -- the time of entry into an industry affects not only their ability to confront uncertainty, but their incentive. The more structured firm, like the timid driver, has an incentive to reduce the amount of uncertainty in its environment, while the more dynamic firm has an incentive to increase its share of the market at the expense of its more cautious rivals. How are these conflicts of interest settled? One firm may control the industry -- as the Ford Motor Company dominated the automobile industry for a few years. Or the older firms in an industry may find some way to prevail by establishing a peaceful arrangement to share markets, which could result in some form of industry-operated cartel (whether official or unofficial) or a government regulated cartel. Which circumstances would be more favorable for self-regulation, and which for government regulation? When entry into an industry is expensive (e.g., as in the automobile or steel industries) self-regulation in the form of price leadership is more likely to occur. But, when entry is relatively inexpensive (e.g., as in the railroads during the nineteenth century or the airline industry at the present time) government regulation is more likely.

As was suggested previously, there can also be differences in interests among consumers. For example, during the 1880s, grain farmers shipping from Chicago to New York probably paid lower rates than those shipping from Fargo to Minneapolis. And I would guess that those farmers who fared most poorly with respect to rates were the most enthusiastic supporters of railroad regulation.

When interests of both consumer and producer groups differ (e.g., in the case of railroad regulation), it can be predicted that both groups will attempt to influence the legislative process. However, regulation can also occur because of different interests of consumer groups (e.g., air pollution) or because the more insecure firms in an industry are able to prevail (e.g., broadcasting regulation).

Behavior of Regulated Agencies

Many political scientists and economists who have worked in the field of regulation are convinced that sooner or later the regulatory agency will be captured by the regulated industry, although there are differences among them with respect to the timing of capture. For example, according to Stigler's theory, regulatory agencies seem to be captured even before they are born. On the other hand, according to Marver Bernstien's observations of regulatory behavior,[13] regulatory agencies go through an entire life cycle before ultimately becoming captured: gestation, youth, maturity, and old age. "Youth" is a period of great uncertainty and conflict between the regulatory agency and the special interests. "Maturity" is a period in which the agency loses its original political support and relies upon standard operating procedures for resolving conflicts. Finally, "old age" sets in when favoritism toward the regulated industry becomes institutionalized.

How can this seemingly irreversible process be explained? According to dynamic thinking, all public organizations are regarded as monopolies in restraint of change. They are so considered because, except during periods of grave national crisis, the hidden foot plays close to a zero role. And, this being the case, it can be safely assumed that in normal times bureaucrats interested in maximizing their longevity will be risk averters. Contrary to common widsom, highly inflexible bureaucratic rules, which have the effect of making bureaucratic change extremely slow and expensive, are not imposed upon bureaucrats by outsiders. Rather, such rules, promulgated by insiders, perform an important role in protecting individual bureaucrats from risk-taking and their organizations from feedback. To be sure, not all public organizations have the same incentive to protect themselves from feedback; generally speaking, those who are forced to live in the most hostile environments develop the most protective mechanisms (for example, public school organizations, the State and Defense Departments). However, prior to World War II, very few, if any, federal government agencies had as great an incentive to protect themselves from feedback. And, if my argument is correct, regulatory agencies do not cease to perform a public service because they are captured by regulated industries. Rather, they make good (but not necessarily ideal) natural partners, because both regulated industries and regulatory agencies are monopolists in restraint of change.

How else might the behavior of regulatory agencies be explained? According to another more static hypothesis, the voter is ultimately responsible. When voters support a reform movement, regulatory agences will act in the public interest. But, when broad political support dies out the incentives of the politicians will change to favor a narrower constituency, namely, the regulated industry.[14] That political factors play an important role in explaining

the behavior of public organizations, no one will deny. If feedback from the political process could have absolutely no impact on regulatory agencies their decline would be fully irreversible. But, it must be observed that to some extent, at least, the process is reversible. For example, the FCC did not permanently block the entry of cable television stations -- by making entry depend upon political support it only slowed the process. Or, to take another case, after a public exposé by Nader's Raiders, the incredibly slow history the Federal Trade Commission was making was temporarily accelerated.

If only an acute political crisis can energize a mature regulatory commission, why were the regulatory agencies initially set up as independent agencies responsible only to Congress? The commonly given reason was the need to remove the regulatory function from politics! According to the first main argument made by the Progressives during the early part of this century, the regulatory commissions were to be flexible and expert in a task often called "scientific decisionmaking" -- a task the British regulatory commissions were alleged to have already mastered.[15] And according to the Progressive's second main argument, the commissions would "protect the rights of those whose resources otherwise left them defenseless."[16] What these armchair reformers left out of account was the all-important matter of feedback: unless a substitute could be found for Jefferson's concept of a modest revolution every twenty years, the commissions would sooner or later become monopolies in restraint of change. However, naive as they were, the Progressives seemingly assumed that the crises which led to regulation could be more or less permanently institutionalized in the regulatory commissions.

Now, if some way could have been found to keep the regulatory commissions in a permanent state of crisis, they would, no doubt, have been entirely different organizations. But, given the fact that the crises did not continue, what constitutes rational behavior from the point of view of the regulatory commissions? When the Interstate Commerce Commission was first established it had to exist in a hostile environment. On the one hand, there was pressure against its involvement in politics. On the other hand, there was pressure from lawyers for the ICC to protect private property from attacks by the majority. If an agency hopes to survive in a hostile environment, what must it do? Generally speaking, best it can, the agency must try to isolate itself from negative feedback. In part, this means denying to the public at large information which might be used to criticize the agency. But, because total isolation is impossible in the real world, it also means heeding only that feedback which might result in serious budget reductions. It is not altogether surprising, therefore, that under the astute chairmanship of Judge Thomas C. Cooley the ICC quickly became a quasi-judicial body. Nor is it surprising that other regulatory agencies were patterned after the ICC. If conservation of the ability to respond to negative feedback is the name of the game, what better an institutional framework than one which simultaneously limits the amount of negative feedback

an agency is likely to receive and insures that the feedback most heeded will come from producer, rather than consumer, groups? Quite obviously, it is feedback from organized producer groups rather than from unorganized consumer groups which determines the budgets of regulatory agencies.[17]

I do not argue, of course, that the political factors are of negligible importance in the evolution of regulatory agencies. The presence or absence of broad political support obviously can influence the speed with which bureaucratic maturity is attained. For example, the loss of political support certainly plays a role in explaining how rapidly the Environmental Protection Agency became a fully mature regulatory agency. However, it would be entirely wrong to assume that were it not for the loss of political support regulatory agencies would be entirely different organizations: they would not, because the internal incentives of both regulated firms and regulatory agencies strongly favor a stable environment.

It should not be assumed, however, that from the standpoint of the regulated industry such a partnership is necessarily an ideal one. Consider, again, railroad regulation. Now, the older railroads had an incentive to obtain a more stable environment, that is, one in which market shares would remain more or less constant. As Paul MacAvoy,[18] Lance Davis and Douglas North[19] have pointed out, from 1871 to 1885, the major railroads engaged in repeated attempts to form a cartel, which more often than not broke down when a new railroad company reduced its rates to take advantage of the elastic demand for long-haul traffic. This action usually resulted in a "rate war" with rates falling as much as 20 percent! As compared with later experiences of the meat packing or automobile industries, this was not war, it was peace. And the customary behavior aimed at avoiding even such minor skirmishes certainly indicates that the railroad companies had a real desire to maintain a stable environment.

On the other hand, from my point of view, it is not necessary to argue that by favoring railroad regulation to the extent that some of the largest railroad companies did,[20] the railroads were able to obtain a near perfect method of cartel management. Indeed, on this score I am more persuaded by the argument of Tom Ulen than I am by that made by MacAvoy, Davis and North. According to Ulen, what made the cartel work as well as it did after the Panic of 1893 was not that the ICC operated the cartel in a different manner than the railroads, but, rather, influence of the merger movement: a network of seven systems emerged which accounted for about 85 percent of railroad earnings.[21] And Ulen argues that because the merger movement resulted in less cheating, the railroad companies really made a mistake in opting for federal regulation. However, from my point of view, the fact that some of the railroads clamored for regulation when it was not needed, only indicates to what lengths they were willing to go to obtain a more stable environment.

As for the ICC, no one will argue that it has not always wanted a stable environment. In fact, it can be argued that the ICC has become so enamoured with protecting its regulations that it has

failed to do a good job protecting the stability of the industry which allegedly captured it. For example, long after the railroads had experienced serious inroads in their business due to competition with trucking, the ICC continued a form of rate discrimination that benefited trucking companies more than the railroads.

The Dynamic Impact of Regulation

This section will provide a simple model of regulation that will help the reader visualize its impact on actual behavior. The argument will begin with a case that involves making substitutions under conditions of weak or statistical uncertainties. Suppose the activities involved in running a household are split into various management responsibilities, including management of the food budget. In the past this responsibility has been exercised mainly by women, but it could be headed by either men or women. Assume further that the managers of the food budget are faced with the following situation: whenever a rise in the price of food occurs the other household members grant them a cost-of-living increase in their food budgets. Of course, the managers do not know on which days particular grocery stores are likely to provide the best bargains. Their objectives will be, first, to go as far as possible to provide their households with a satisfactory eating experience, given their budget constraints; and, second, each manager will try to put aside some pin money for a rainy day.

How should they go about satisfying both of these objectives simultaneously? Obviously, if the managers go too far by way of pleasing the food preferences of each household member, they will not be able to satisfy their budget constraints nor put aside any pin money. On the other hand, if they are single-minded about concentrating on building up their stock of pin money at a maximum rate, the managers of the food budgets will risk having a revolution on their hands. So, to avoid both of these dangers, feedback in the form of either family protests or a budget crisis must be heeded. In short, although the risks are different, food budget managers, like business entrepreneurs, must constantly weigh two risks.

How can the various food budget managers go about their tasks in a way so as to maximize their dynamic efficiency? They must cooperate by competing. Suppose that some are more skilled in discovering bargains and others are more skilled in the art of cooking. Then the way to make best use of their respective talents is to offer a system of prizes that will encourage the discoverers to compete to locate the best bargains, and will allow the other managers to act as borrowers. If such competition occurred, no one could predict beforehand just which bargains would be found on a particular day. Indeed, it is the very inability to make predictions on the basis of initial conditions that allows the food managers to do as good a job as they can in minimizing their two risks.

Now, let us suppose that the less successful discoverers would like to modify the rules of the game to enhance the probability of their own survival. And, furthermore, imagine that a group of less successful discoverers is united with a group of consumers who demand a greater degree of certainty in their environments with respect to the food they are served from day to day. If this coali-

tion results in some form of regulation which succeeds in reducing the degree of uncertainty with respect to food purhases, the incentives of entrepreneurs to engage in dynamic behavior and their ability to act in a dynamically efficient manner will be affected.

The Incentive Effect: As we already have seen (Proposition III, Section I), the greater the degree of certainty introduced into an environment, the smaller will be the incentive to engage in dynamic behavior. If, in the above example, regulation has the effect of introducing more certainty into the environment, the incentives of all food budget managers will be changed. In the limiting case -- 100 percent regulation and a zero degree of uncertainty in the items to be served from day to day -- the food managers would have no incentive whatsoever to engage in dynamic behavior. And, while in a world of strong uncertainties regulation can never completely succeed in eliminati g uncertainty in a particular industry, the more it does succeed, the more it will have a negative impact on incentives.

The Uncertainty Effect: In addition to the incentive effect, regulation has a negative impact on dynamic efficiency, because it jeopardizes the ability of an entrepreneur to take full advantage of the uncertainties confronting him. Suppose, for example, that the regulatory authorities order the food managers to serve chicken every Sunday. What affect will this have on their ability to deal with uncertainty? To be sure, they might be lucky in locating a supermarket where bargains were to be found on Saturdays. But, generally speaking, the greater the degree of constraint imposed on entrepreneurs' ability to deal with uncertainty, the smaller the chance of their living within a fixed budget constraint, inasmuch as the greater the degree of constraint the more outcomes are dependent on luck and luck alone.

To be sure, from the perspective of static economics, uncertainty is an annoyance which serves only to complicate the entrepreneur's calculations. However, from the perspective of a dynamic world -- a world of imperfect knowledge -- uncertainty is to be regarded as a valuable asset. For example, by manufacturing automobiles rather than going into the pin making business, Henry Ford went into a field in which the uncertainties were several orders of magnitude greater. But it was precisely this greater degree of uncertainty that enabled him to make several orders of magnitude more money than he otherwise might have.

Regulation protects some firms from creative destruction at the expense of robbing others of a valuable asset: their opportunities for making good use of uncertainty. For example, suppose that the manufacturers of steam cars had managed in, say, 1895, to establish an equivalent of the ICC in the automobile industry. What would the effect have been? Quite obviously, while preservation of the status quo would have protected the property rights of steam car manufacturers, it would only have done so at the expense of depriving those like Henry Ford of an important freedom: the freedom to make the best possible use of uncertainty.

The constraints imposed on the entrepreneur's freedom of choice for dealing with uncertainty may take on a variety of specific forms. Nevertheless, we can think of an uncertainty trade-off showing that, while regulation can make some outcomes more predictable, it

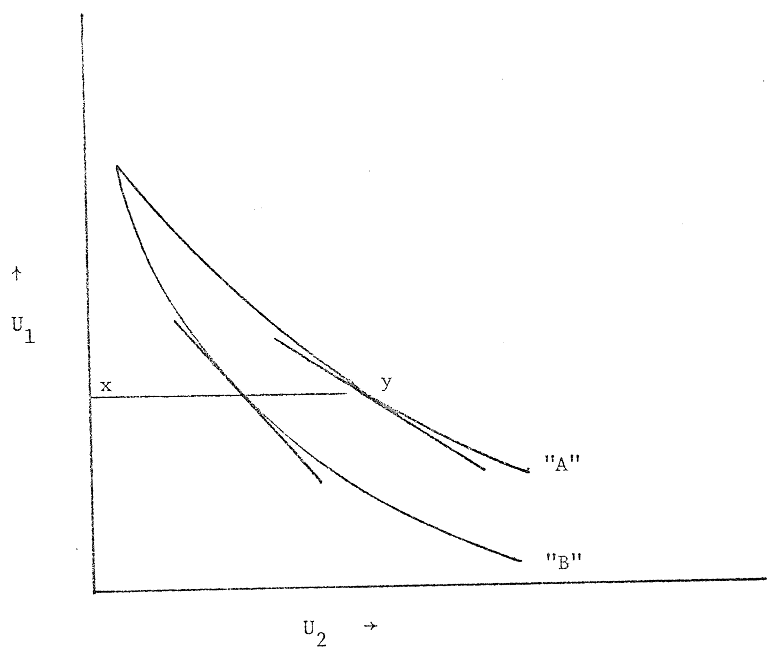
only does so at the expense of making other outcomes more unpredictable. In the above example, we can postulate a trade-off as is shown in Figure 3, in which U_1 is the degree of uncertainty with respect to the household environment (as measured in day-to-day food purchases), and U_2 is the degree of uncertainty with respect to the probability of keeping within the food budget. Even if we assume a world of weak or statistical uncertainties, it is impossible to introduce more certainty into the household environment by applying a higher degree of regulatory constraint without introducing more uncertainty into the probability of keeping within the budget constraint. The greater the degree of uncertainty, the greater will be the cost of applying a particular degree of constraint. If it becomes impossible to increase the budgets of the household managers in the same degree that the cost of living goes up, the managers will be faced with a higher degree of uncertainty, which is indicated by the more rapidly falling curve "B." Under these circumstances, a given degree of constraint (e.g., x,y) will cost more in terms of the entrepreneur's freedom of choice for dealing with uncertainty.

In short, the economic effects of regulation are, first, to dull the incentives for dynamic behavior; and, second, to jeopardize the attainment of dynamic efficiency. However, it should not be assumed that these effects are peculiar only to government regulation. As we already have seen, whenever incentives for dynamic behavior are weak, it can be predicted that some type of regulation will emerge, whether by the dominant firms in an industry or the government.

To be sure, not all economic regulation is equally bad. Generally speaking, regulation is likely to achieve its poorest

FIGURE III

THE UNCERTAINTY TRADE-OFF



results when regulatory agencies act as if their current information represented permanent truths, and its best results when an incentive is provided for reducing costs. For example, it has been argued that the two or three year delay in reducing long distance and international telephone rates in response to reductions in costs supplied an incentive for reductions in costs similar to that provided by competition. And, although the reductions of costs and rates were not nearly as spectacular, the same argument can be applied to public utility rate-making. However, as Paul Joskow has shown in his studies of public utility rate-making, the standard operating procedures of public utility commissions make action to change rates more likely when costs are rising than when they are falling.[22] Hence, when costs are rising, pressure from public utility companies can transform a contract that provides for a sharing of the benefits of cost reductions with subscribers into a cost-plus contract.

In other words, at its very best, regulation can only temporarily provide the incentives associated with dynamic competition. And the more typical consequence is to both dull the incentives to engage in dynamic behavior and constrain the opportunities for dynamic efficiency.

What tests would enable us to decide whether dynamic theory makes good or bad predictions? One test can be related to those who succeed to top-management positions after regulation is adopted. For example, it is my prediction that after the steel cartel was formed the people who succeeded to top-management positions were better at manipulating the environment to suit the technology than manipulating the technology to suit the environment. Thus, when rapid progress is made in reducing the cost of steel, we should expect to find people like Andrew Carnegie a fairly common type of leader. On the other hand, after the cartel was formed to help the steel industry make slow history we should expect to observe the top-leadership posts going to people like Judge Gary -- if ever he saw a blast furnace, it probably was not until after he was dead. And I would make exactly the same prediction for the kind of person likely to become a big wheel in a railroad company, before and after the beginning of cartelization and regulation.

III. The Impact of Economic Regulation

The first step in estimating the dynamic costs of regulation consists of describing its impact in qualitative terms. In this section I will indicate, by way of example, that the incentive and uncertainty effects of regulation not only occur, but can have an important impact. The examples fall under two general headings: regulations concerned with characteristics of the final product; and regulations concerned with minimum rates.

REGULATIONS CONCERNED WITH THE CHARACTERISTICS OF THE PRODUCT

Several important regulatory activities come under this heading, including regulation to reduce gasoline consumption, environmental regulations, and regulations imposed by the Defense Department on weapon systems development. It will be seen that such regulation has the effect of reducing feedback, which makes preferred outcomes

from a regulatory point of view more certain at the expense of making dynamic efficiency (whether measured in terms of reductions in costs or improvements in quality) more uncertain -- with a narrower diversity of ideas, progress will depend on luck and luck alone. It also will be seen that the importance of the impact, in terms of narrowing the diversity of ideas, will depend upon the degree of constraint involved.

Reducing Gasoline Consumption: Consider the form of regulation now employed to achieve reductions in automobile fuel consumption. Automobile firms are required to bring about by specific dates specified reductions in the average fuel consumption of all the various automobiles produced. Now, it is obvious that this form of regulation permits better use of uncertainty than one which might seek to impose scheduled reductions in fuel consumption on each model of car: when feedback indicates that reductions in fuel consumption are easier in some models than in others, manufacturers can concentrate their efforts accordingly.

On the other hand, as compared with a form of regulation that provides better incentives for dynamic behavior, the current method of regulation does not result in a wide search for new ideas. Consider, for example, a gasoline tax high enough to make the elasticity of demand for automobiles highly dependent on their fuel consumption. With such a tax, incentives would favor rapid progress in reducing fuel consumption. Indeed, it can be said that the current method of regulation provides perverse incentives for generating a diversity of ideas: if an entire industry concentrates on the same technical approach it is easier to get the regulations relaxed.

On the other hand, from the point of view of the automobile industry, direct regulation is preferable to incentive regulation: if it is assumed that the name of the game is sharing markets, while trying to keep imports to some minimal figure, then, in terms of these objectives, the current method of regulation is preferable because it results in less feedback and more predictable outcomes.

Regulation to reduce automobile emissions does not permit automobile manufacturers the degree of freedom associated with regulation to reduce gasoline consumption. But as compared with a tax to reduce emissions, the impact of regulation is the same: it will result in the generation of a narrower diversity of ideas. But, again, regulation is likely to be favored over a tax, because it makes market shares more predictable.

Regulation of Military Weapon Systems: No regulatory agency pursues a policy of applying constraints more thoroughly than doe- the Defense Department. As a means of banishing uncertainty elaborate paper studies are undertaken. For example, in the C5A program, 35 tons of documentation were delivered to the DOD at a cost of 60 million dollars.[23] And on the basis of these studies elaborate technical and operational requirements were prescribed that left little or no opportunity for those engaged in development to use their imaginations.

What can be said about the cost to the taxpayers of such regulation? Let us consider two more or less similar airplanes: the C5A developed by Lockheed for the Air Force and the Boeing 747

developed for commercial airlines. What role did the requirements process play in making the former twice as expensive to produce as the latter?[24] As it happens, the C5A had to be optimized for shorter runways and lighter gross weight, which led to an uphill battle to save weight and a skyrocketing of production costs. In short, the story is the familiar one of obtaining a modest gain in performance by doubling the cost. Though feedback during the development program certainly suggested that the weight-saving program would be very difficult, the highly bureaucratic process employed in the management of weapon systems was insensitive to this information.

Moreover, it is by no means clear that the decisions would have been any different had the final cost of the C5A been known from the outset. The military officers in charge of weapons programs are not promoted on the basis of meeting a cost target, rather, important to their careers is developing the weapon according to a fixed schedule and a tight initial requirement. Furthermore, firms engaged in military weapon system development projects have no particular incentive to minimize costs: the profits made in relation to sales is strictly limited, and any individual program faces a highly inelastic demand with respect to price, because if an overrun occurs the adjustment is commonly spread among a number of programs. So, while there is an incentive during the development of commercial airliners to heed feedback suggesting that demanding the last increment in performance will double costs, no such incentive exists in the development of military systems.

Indeed, the military and commercial markets are so different from one another that firms engaged in both defense and commercial work commonly segregate their military from their nonmilitary projects.

As another illustration, consider two competing air-to-air missile development projects: the Sidewinder and the Falcon. The Sidewinder provided quite as good an operational capability as the Falcon, however, its development cost was only about 10 percent as great, and its procurement cost, about 15 percent as great.[25] How are these differences to be explained? As it happens, the infrared technology upon which the Sidewinder is based makes possible a missile which is simpler to develop and easier to produce. However, it was not luck and luck alone which resulted in the choice of a missile based upon infrared technology. The Sidewinder was developed by a naval laboratory whose civilian leadership adopted as its mission the development of a low cost and reliable system. This laboratory wisely sponsored internal competition prior to the decision on the Sidewinder. In short, whereas the Sidewinder was developed from beginning to end by making good use of uncertainty by defining the "system" very broadly and taking feedback into account, this procedure was not possible with respect to the Falcon program.

Why was there a difference of a factor of five between the costs of the Sidewinder and Falcon programs, and only a difference of two between the 747 and C5A programs? Undoubtedly more uncertainty was involved in the former than in the latter programs. And, as was pointed out in Proposition III, Section I, with more uncertainty involved the cost of applying more or less equal degrees of constraint is much greater.

To be sure, on the basis of only two cases definite conclusions cannot be arrived at as to what it costs the Defense Department in its research and development activities to minimize bureaucratic risk. Nevertheless, it is my contention that a systematic bias in internal incentives occurs which pushes the DOD to undertake too many programs conducted as was the Falcon, and too few like the Sidewinder. This bias results from bureaucratic incentives which favor keeping within a rigid schedule and narrow technical constraints rather than a tight cost constraint. How might the internal incentives be changed? Suppose that the Secretary of Defense promulgated a policy which resulted in a high probability of program cancellation whenever actual costs exceeded the initial target by, say, more than 10 percent. Were such a policy adopted, the environment in which weapon systems are developed would be made more uncertain; and with more uncertainty in the environment there would be a greater incentive to take risks to minimize cost escalation.

<u>FDA Regulations</u>: A form of regulation whose impact is more difficult to judge is that which was undertaken by the Food and Drug Administration (FDA) in the early 1960s. It resulted from complaints brought before Senator Kefauver's Antitrust and Monopoly Subcommittee, namely, that through advertising ethical drug firms were concocting drugs with questionable advantages and making them seem like wonder drugs. Therefore, as a consequence, new regulations were enacted to require a testing procedure for determining not only the safety but also the effectiveness of new drugs.[26]

What can be said about the impact of regulation upon R & D costs of new drugs? Several studies have suggested that regulation was the cause of a two-to-threefold increase in development times and costs.[27] Even before the new regulations were adopted, however, there was a large decline in new discoveries: from 1958 -- the highpoint -- until 1962 -- when the new regulations were adopted -- the number of new chemical entities introduced annually declined by more than 50 percent.[28] And almost invariably when opportunities in a particular field dry up, there is an exponential increase in the cost of making progress.

A better argument with respect to the impact of new regulations on costs was made by Henry G. Grabowski and John M. Vernon, who compared United States and United Kingdom discoveries per dollar of R & D investment before and after regulation. They found that regulation roughly doubled R & D costs.[29] Nevertheless, their argument is not convincing, because no mechanism was provided to show how the new testing procedure either affects the incentives of firms or their ability to deal with uncertainty. In the case of military weapon system development quite specific constraints are imposed before development is begun. But there is no reasons to suppose that, in itself, testing will result in a large increase in R & D outlays. Indeed, if the Defense Department substituted a rigorous testing procedure for its requirements process, R & D and production costs would be much smaller than they are today.

This is not to say that the testing procedure could not result in the generation of a narrower diversity of ideas: drug firms could become so preoccupied with making a good showing on the tests

that they might fail to make good use of uncertainty. However, more would have to be known before it could be ascertained that such an effect has occurred.

REGULATION CONCERNED WITH MINIMUM RATES

There are regulations aimed at minimizing -- if not precluding -- rate cutting: ICC regulation or railways, trucks, and waterways comes in this category, as does FAA regulation of the airlines. As in the case of product regulation, the impact of such regulation is to narrow the diversity of ideas so as to make lower cost outcomes more uncertain, inasmuch as such outcomes are more dependent on luck. For example, jet planes not only provide a faster, but, as became evident only in retrospect, a cheaper form of transportation than piston-driven airplanes (with the 707 jetliner, operating costs, including depreciation, were only half those of the DC-3).[30] Or, to take another case, due to inventions like piggybacking, unit trains, and Big John Cars, from 1948-1963, productivity in the railroad industry rose 3.4 percent annually, almost twice the average rate during the period 1919-1937.[31] The reason for this impressive spurt in dynamic performance, no doubt, was a large increase in the supply of negative feedback, when due to the artificial advantage provided by value-of-service pricing railroads found that while their shipments were declining, those of trucks were increasing.

On the other hand, the probability that such lucky accidents will benefit a regulated industry over a period of, say, 50 years, is something like one in a million. The reason: an environment featuring little risk-taking does not provide an incentive to hold wages and prices in check; rather, such an environment favors monopoly costs. For example, it is no accident that featherbedding practices thrived in their full blown form first in railroads and more recently in trucking. Nor is it any accident that some airline companies have displayed the same degree of bureaucratic behavior as did the railroads in the early 1900s, or that in the airline and trucking industries "service" competition has tended to dominate rate competition. These are mere indicators of an environment that, due to the absence of the hidden foot, features monopoly costs.

To be sure, even with the complete absence of rivalry there would be some incentive to reduce costs, because it is only through cost-saving inventions that the trend toward monopoly costs can be slowed. However, technological progress which does not come about as a result of rivalry between firms is unlikely to have significant benefits for the consumer.

IV. The Longer-Run Costs of Regulation

We have seen that the incentive and uncertainty effects of regulation are important. And in this section I will argue that because they are important effects, generally speaking, the dynamic costs of regulation are likely to exceed the static costs.

Example 1: Regulation of Railways and Trucking: What can be said about the relative costs of regulation as measured in static and dynamic terms? As far as railways and trucking are concerned, it is generally agreed that the principal static efficiency loss results from the fact

that much traffic sent by truck for distances of 200 miles or longer could be more cheaply carried by railroads. According to Ann Friedlander, the total welfare loss resulting from value-of-service pricing and other misallocation of resource costs is in the neighborhood of 500 million dollars annually[32] -- some four percent of the total costs of regulated motor and rail carriers.

By contrast, dynamic theory assumes that while at any moment some forms of transportation will have a relative advantage in particular applications, it is competition for the "marginal" customers (those whose preference for one mode over another is relatively slight) which results in rivalry for markets and a keen incentive to reduce costs. How might we go about making quantitative estimates of the benefits? To do this we must be guided by historical data.

We do know that in the railroad industry a relatively long time is required for inventions to become generally adopted: for the diesel locomotive the period was 15 years before adoption by 90 percent of the railways, for the Mikado locomotive, 25 years, for four-wheel trucking locomotives, 20 years, for centralized traffic control, 25 years, 30 years for retarders, and over 30 years for piggyback and unit operations.[33] And we can assume that the primary reason for this low adoption rate was an environment which featured little risk-taking. In industries which feature more risk-taking, the adoption rate for inventions of the same significance for reducing costs is commonly more than twice as rapid.[34]

If the result of more rivalry is a significant shortening in the period of adoption, by how much might the rate of total factor productivity (increase in output divided by the weighted inputs, including capital) be increased? As already was noted, we do know that in the period after 1948, when several important labor-saving devices were adopted in quick succession, the rate of total factor productivity advance in railroads almost doubled. To be sure, there was no such increase in productivity gains in trucking; but, still, the weighted average for trucks and railroads increased from about 2.4 percent to 4.0 percent annually (or by more than 50 percent).[35] Such an increase seems to provide a reasonable estimate of the benefits derived from deregulation. Actually, this is a conservative estimate inasmuch as it assumes that a more challenging environment merely speeds up the diffusion of technology -- but does not make for a larger pipeline of inventions.

How do these gains compare with those from better resource utilization? In little more than three years the dynamic gains would exceed the static gains; and in ten years they would be about three times as large. Admittedly, the estimated of the dynamic costs are subject to a much larger margin of error. Nevertheless, it can be safely assumed that over a period of several decades the dynamic costs will dominate the static costs.

Example 2: Airline Regulation: It is generally recognized by economists that the main benefit from airline deregulation will be rate competition in addition to the service competition in which they are now engaged.

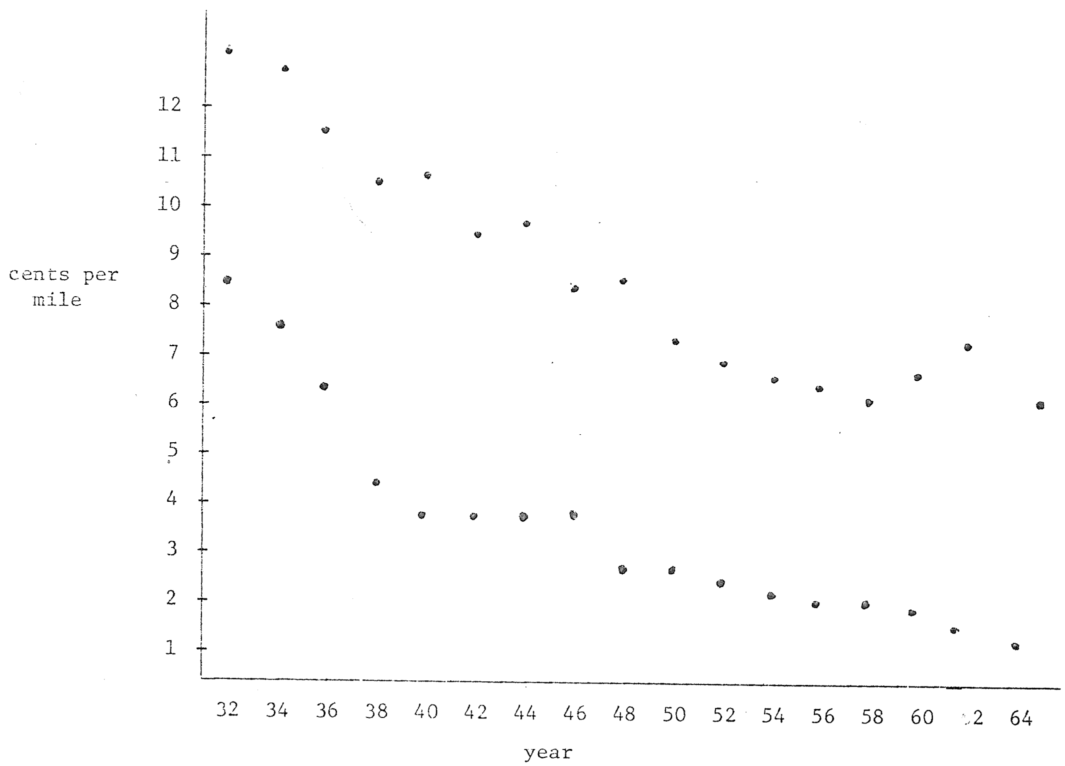
And it is no small wonder that price competition has been equated with easier entry to the industry. Whether the new airline is an intra-state airline, like PSA, or an international airline, like Lakers, generally speaking, new airlines have played the same role in initiating rate competition as did relatively new railway companies during the period 1870 to 1890.

What can be said about the relationship between airline rates and the operating costs of airliners, including depreciation? Figure IV shows average rates during the period 1932 to 1964 and the average cost of the three most economical airliners developed during the preceeding three years (with both rates and costs measured in constant dollars). The reduction in costs, it can be seen, was fairly spectacular: from 1932 to 1948 per seat mile costs were reduced by two-thirds, and from 1948 to 1964, by about half. For the period 1948 to 1964, these reductions are in close agreement with John Kendrick's estimates on total factor productivity for air transportation (according to his figures, during the peiod 1948 to 1964, it more than doubled).[36]

The reduction in rates was not, however, so spectacular. During the period 1932 to 1948, they declined by about half as much as costs, and during the period 1948 to 1969 about one-third as much. (Figure IV) The failure of higher costs to decline more rapidly or airline profits to increase significantly during the latter period is entirely understandable in terms of the tendency for regulation to feature monopoly costs. More surprising is the failure of rates to decline more rapidly during the 1930s. However, the only period during which there was real pressure on the aircraft companies to develop more economical airliners was

FIGURE 4

REVENUE PER PASSENGER MILE AND AIRCRAFT OPERATING COSTS
(Including Depreciation)
1932 - 1964
(Constant Dollars)



Costs computed on basis of three most economical airliners available during three preceeding years.

Source: "Air Transportation in the United States," pp. 132,133, & 135.

that between 1934 and 1938. Before 1934, joint ownership of airlines and aircraft companies was common, and because a very significant proportion of revenues came from the airmail subsidy, it was the practice of airline companies to feature speedy airliners initially developed for the military services. It is no accident, therefore, that beginning in 1935 a series of economical airliners began to appear: the DC-2 (6.8 cents per seat mile), the Lockheed L-10 (4.7 cents), the DC-3 (3.3 cents), and the Boeing 307 (3.2 cents).[37]

What is the prognosis for the future? If the airline industry is not deregulated with respect to rates, and if entry is again substantially closed, rates will not continue to decline. Rather, with no competitive pressure to reduce costs, and with continued pressure for higher wages, it can be expected that, as measured in real terms, rates will increase at least 2 or 3 percent a year. On the other hand, with deregulation and easy entry, as a minimum it can be expected that rates and costs will decline by as much as they did during the period 1948 to 1964, that is, by 1 or 2 percent annually. To be sure, rapid improvements in the efficiency of airliners no longer can be expected. But no more than improvements in efficiency on the railroads depended entirely on better locomotives, does further improvements in airline efficiency depend upon improved airliners.

How much consideration should be given to the complaints of those railroad or airline firms which would not be favored by deregulation? For obvious reasons, firms in a relatively poor competitive position worry mainly about their short-term profits. However, it is very questionable whether the longer-term stability of firms in these industries is promoted by regulation. On the contrary, to the extent regulation reduces the degree of uncertainty firms have to live with, it makes them less able to deal with new circumstances -- and, by being less able to deal with new circumstances, they are not as stable from a dynamic point of view. In Britain the commonly given reason for nationalization of the airlines and railroads is that with private ownership it was impossible to obtain adequate funds for modernization. However, the more fundamental reason these industries were finally nationalized may be that regulation almost completely sapped their ability to engage in dynamic behavior.

Example 3: The Current Institutional Framework for Supplying Electric Energy: As matters now stand, the United States will be heavily dependent upon a partnership between a government agency (ERDA) and electric public utility companies for dealing with one major aspect of its future energy problems. If no substantial measures are taken to reform this partnership, what can be said about the prospects for developing reasonably economical forms of energy in the future? As it happens, in the form of developing nuclear energy plants a number of "experiments" have been made which are relevant for answering this question. Their history can be very briefly summarized as follows: First, although in 1953 the Eisenhower Administration decided to press for an early capability in the field of nuclear reactors, and, although Admiral Rickover urged the AEC to emphasize a very conservative design (the pressurized water reactor -- PWR) which already had been used in submarine reactors, in the power reactor demonstration program begun in 1955 the AEC began to subsidize a power reactor demonstration program

which included several different designs: the PWR, the sodium graphite reactor, the boiling water reactor, the homogeneous reactor, and the fast breeder reactor. With only one or two exceptions, however, out of more than two dozen projects undertaken, the only successful ones involved using a proven design. And in 1957, the AEC abandoned a program of multiple approaches in favor of one which would emphasize two designs: the PWR for an early capability and the fast breeder reactor for a later capability.

Second, in late 1962, General Electric and Westinghouse began to offer "turnkey" contracts -- fixed-price contracts under which the builder assumed complete responsibility until a plant was turned over to a utility for operation. The principal difference between the turnkey plants and those built previously is that they were much larger: in the 400 to 800 Mwe range rather than in the 100 to 200 Mwe range. However, the actual scaling problems proved to be far from simple, the actual economies of scale disappointing, and the actual costs roughly twice the initial estimates.[38] So, after losing something like one billion dollars, in 1966 both companies discontinued turnkey contracts.

Third, notwithstanding the history of the turnkey projects, even before their completion public utility companies began to build nuclear plants with no subsidy from the Federal Government. In 1967, 30 reactors were ordered, in 1968, 14, and in 1969, 7. The period 1970 to 1974 also featured rapid expansion. However, although the number of reactors ordered rose to 36 in 1973, by 1976 it had declined to one.

Fourth, the cost history of nuclear reactor plants is highly reminiscent of that of military weapon systems. According to information contained in a study made by David Montgomery and James Quirk, the overrun's of the post-turnkey era were even larger than those of the turnkey era itself: on plants ordered in 1967 and in 1968, the final costs came to between 2-1/2 and 3 times the earlier estimates.[39] Moreover, as more plants were built, the actual deflated costs did not become smaller, they became larger: if the costs of nuclear units had risen no more than the construction price index and the increase in interest costs, then, from 1967 to 1976, the cost of nuclear energy would have roughly doubled. But, because costs increased by a factor of four, nuclear costs per kilowatt were brought well above those of coal.[40]

How is the failure to develop a more economical nuclear energy capability to be explained? From the perspective of static economics, in which all alternatives are regarded as givens, it simply would be assumed that the alternative which emerged from competition -- the PWR -- was in the short-term, at least, the most efficient reactor. Moreover, it would be assumed that cost escallation was the result of extraordinary events over which the developers had no control: material and component shortages, labor difficulties, and the more stringent development procedures adopted in the early 1970s. And there is no doubt that the more stringent procedures did impose serious constraints on the ability of utility companies to deal with uncertainty; in fact, utility engineers may have found themselves dealing with so much uncertainty that they were overwhelmed.

Nevertheless, from the perspective of dynamic theory, asking how special events of one kind or another may have contributed to cost escallation is asking the wrong question. In the real world, programs are always going to be affected by special events. However, as we have seen, organizations strictly optimized in terms of the goal of static efficiency are going to have a negligible ability to be dynamically efficient. How might firms having a negligible ability to deal with uncertainty behave? They undoubtedly would behave as if construction of a nuclear power plant was a straightforward job, involving little or no uncertainty. And based upon evidence taken from a RAND report on the Development and Commercialization of the Light Water Reactor, 1947-1976, not only the utility companies, but Westinghouse and General Electric as well, behaved in the manner just predicted.[41] For example, people expert in dealing with uncertainty might be expected to know that nuclear technology would provide quite as challenging a task in scaling as one might expect to find. Yet, according to the RAND report just cited, in the turnkey programs both General Electric and Westinghouse assumed "that sufficient experience had accumulated in the demonstration projects to support confident estimates of the costs of designing, developing, building, and operating reactor-powered generating plants in the size range between about 400 and 1,000 Mwe."[42] Or, to take another example, an organization expert in dealing with uncertainty, and concerned with minimizing the cost of an entire nuclear power plant, might be expected to recognize that it was unlikely a low cost solution could result from mating a nuclear reactor with an unmodified turbine section of a conventional power plant. Yet, according to the same report it was commonly assumed that "the balance of plant -- that portion of the nuclear power plants exclusive of the nuclear steam supply system -- would in most respects resemble the turbogenerator systems of fossil-fuel plants of similar generating capacity."[43]

Nor, according to the same report, were the AEC scientists really aware of the uncertainties involved in minimizing the economic costs of nuclear reactors:

> The government sponsors were preoccupied, at times almost to the exclusion of other considerations, with finding avenues for demonstrating variant reactor designs and with solving a number of technical problems that troubled design, construction, and operation. There were, indeed, recurrent expressions of concern for the economics of nuclear power, but the subject often was treated as though technology alone would provide adequate means of settling all questions of commercial application.[44]

As it happens, the framers of the 1954 Atomic Energy Act were aware of the dangers which can arise when technologists do not have to think in terms of the entire system and meet tight cost constraints. Indeed, as a Joint Committee background report indicates, appreciation of this danger was the rational for the partnership conceived by the 1954 law:

> In particular, we do not believe that any developmental program carried out solely under governmental auspices, no matter how efficient it may be, can substitute for the cost-cutting and other incentives of free and competitive enterprise.[45]

The problem with this statement, as with the common wisdom upon which it is based, is that private activities differ enormously in the degree of risk entrepreneurs are willing to take, and in the incentives, therefore, for cost-cutting. When it comes to developing low-cost alternatives, there is no substitute for an environment in which a firm's very survival depends on its ingenuity to generate new ideas. In other words, there is no substitute for an environment which contains a hidden foot as well as a hidden hand. For example, we can be reasonably certain that as of 1910 the internal combustion engine represented a good choice, not simply because several major types of engines were concurrently under development, but rather, because there was an appropriate environment to insure entrepreneurs would excercise a high degree of ingenuity in taking full advantage of the inherent potentialities of each of the alternatives. And for the same reason, we can be reasonably sure that no major opportunity was overlooked for making gasoline powered cars as inexpensive as possible.

The regulated utility industry, however, neither possesses the incentives nor the type of R & D organization needed to discover which particular type of reactor is likely to be most successful. What would have happened if commercial airplanes had been developed on the basis of the same kind of partnership? My best guess is that the technology would never have progressed beyond adaptations of bombers developed for commercial use -- planes such as the Boeing 221, which was developed in 1930 and whose operating costs were about six times those of the DC-3.[46] What would have happened if the same type of partnership had been employed not only in the United States, but in other countries as well, to develop automobiles? That particular industry would never have progressed beyond steam cars costing something like 20,000 dollars each. To be sure, in retrospect this might seem like a favorable outcome. But, as was observed earlier, because regulation minimizes the diversity of ideas, only luck and luck alone can result in favorable outcomes.

The reason the electric utility companies cannot be expected to minimize the costs associated with any particular type of reactor is the same reason they are not likely to discover the most promising designs: their organizations are optimized for making tiny improvements in an existing technology which lead to small but steady reductions in energy costs. However, organizations so specialized are no more capable of discovering major possibilities for reducing the costs of a new technology than a probate law firm is capable of trying a criminal law case.

Hence, it is not altogether surprising that the reactor design chosen for the short-run term was the PWR. The PWR might not have been the most economical design; but it certainly was the design best matched the technological capabilities of utility firms. Nor is it surprising that the cost history provides little or no evidence of learning. For learning to have occurred, the organizations in question would have had to have some capability for engaging in dynamic behavior.

This is not to say that partnerships between government and industry cannot work. In the case of commercial aircraft, the partnership between NACA and the aircraft companies worked splendidly. However, in that case there was keen rivalry between aircraft firms -- rivalry which resulted in a demand for new ideas and organizations optimized to make good use of new knowledge. But not only is such rivalry absent in public utility industries, as H.S. Burness, D. Montgomery and J. Qu rk have shown, given the way utility rates have been made over the past fifteen years, utility companies have an incentive to transmit the risks of cost escalation to their customers.[47] Associate a dynamic public laboratory with a static industry possessing such incentives, and not only will the laboratory lose its dynamism, but its role will change from a generator of a wide diversity of ideas to an advocate of a few pet ideas.

How might this danger be avoided in the future? Quite obviously, the electric utility industry must be restructured so there can be more rivalry in generating activities: various generating activities (e.g., coal solar, nuclear energy, etc.) must be placed in the hands of independent firms, thereby separating generating activities from transmission and distribution activities. Such a restructuring would not only make entry into the industry easier, and provide better incentives, it would also make possible a type of scale economy that vertically integrated firms are unlikely to enjoy: a scale economy resulting from the ability of a single firm to benefit from experience in building a number of reactors. Two or three nuclear power plant firms are more likely to benefit from experience in finding ways to minimize costs than dozens of public utilities who look upon each installation as a custom-design job.

Would the gains in dynamic efficiency exceed the losses in static efficiency? Although this is a very complex issue that will require more research, it seems unlikely that genuine economies of scale occur beyond the plant level. To be sure, utility companies perform a function in combining different types of plants to make it possible to meet peak loads in a reasonably efficient manner. However, there is no reason to suppose that this function cannot be performed by an appropriate pricing scheme. In fact, Derek McKay has argued that with such a scheme it would be possible to bring about a gain in static efficiency.[48]

The real question in deciding how important it is to restructure the industry is how to define an "energy shortage." If a shortage is defined as the ability of the United States to supply larger and larger amounts of energy in the future (though at a less rapid rate of increase than occurred in the past), then a convincing argument probably cannot be made for restructuring the utility industries, because this objective probably can be realized by relying mainly upon coal and solar energy developed to be sold directly to the final consumer, rather than to be delivered through existing transmission facilities. To be sure, clean coal might turn out to be more expensive than the coal we have today, and the ultimate potential for economical solar energy probably would not be fully realized. But, if a "shortage" is defined from an engineering point of view, cost is not an important consideration. On the other hand, if a shortage is defined as the prospect of much higher energy costs than would be necessary with a greater dynamic capability, then the argument for restructuring

the industry becomes very strong: the longer-run costs easily could be to make electric energy twice as expensive as need be.

In conclusion: there is no guarantee that deregulation will in itself result in an environment featuring a significantly higher degree of risk-taking. When it does not, achievement of real gains in dynamic efficiency will require a restructuring of the industry. However, not only are the longer-run economic costs of regulation very high, but so are the political costs. A society which encourages to the fullest extent possible the generation of a wide diversity of ideas provides the best safeguard against despotism. However, because regulation is incompatible with such a society its longer-run political costs cannot be ignored.

FOOTNOTES

1. H. Averch and L. Johnson, "Behavior of the Firm Under Regulatory Constraint," American Economic Review 52 (December 1962).
2. George J. Stigler, "The Theory of Economic Regulation," Bell Journal of Economics 2 (Spring 1971):3-21.
3. For further discussion, see Burton H. Klein, Dynamic Economics, (Cambridge: Harvard University Press, 1977), Chapters II and III.
4. Richard P. Feynman, Robert B. Leighton and Matthew Sands, The Feynman Lectures on Physics (Reading : Addison-Wesley Publishing Company, 1966):44-2,44-10.
5. Ibid., pp. 37-8 ff.
6. Axel Leijonhufuud, On Keynesian Economics and the Economics of Keynes, (London: Oxford University Press, 1968):70-5.
7. Kenneth J. Arrow, "Uncertainty and the Welfare Economics of Medical Care," American Economic Review 3 (December 1963): 941.
8. For further discussion, see Dynamic Economics, pp. 35-52.
9. Ibid.,
10. Ibid., pp. 100-101.
11. John E. Tilton, International Diffusion of Technology: The Case of Semiconductors (Washington, D.C.: Brookings, 1971).
12. Dynamic Economics, P. 17.

13. Marver H. Bernstien, Regulating Business by Independent Commission, (Princeton: Princeton University Press, 1955).
14. Barry R. Weingast, "A Representative Legislature and Regulatory Agency Capture," PhD Thesis, California Institute of Technology, September 1977.
15. Regulating Business by Independent Commission, pp. 24-5.
16. Ibid.
17. For a similar argument regarding the behavior of politicians who want to maximize their longevity, see Morris P. Fiorina, Congress: Keystone of the Washington Establishment (New Haven: Yale University Press, 1977), Chapter 7 and 8.
18. Paul W. MacAvoy, The Economic Effects of Regulation, (Cambridge: MIT Press, 1965), Chapters 3 and 4.
19. Lance E. Davis and Douglas C. North, Institutional Change and American Economic Growth, (Cambridge University Press, 1971): 136-67.
20. For an account of the role of the railroad companies in the establishment of regulation see Gabriel Kolko, Railroads and Regulation, 1877-1916, (New York : W.W. Norton & Co., 1965):30-45.
21. Tom Ulen, "The Interstate Commerce Commission as a Cartel Manager: Was It Necessary?", unpublished paper, Stanford University, March 1977.
22. Paul Joskow, "Inflation and Environmental Concern: Structural Change in the Process of Public Utility Price Regulation," Journal of Law and Economics 17, (October 1974):291-328.

23. Adam Yarmolinski, The Military Establishment, (San Francisco: Harper & Row, 1971):272.

24. Ibid., p. 273.

25. Ibid.

26. For further discussion see Sam Peltzman, "An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments," Journal of Political Economy 83, (1975):1049-91.

27. V.A. Mund, "The Return on Investment of the Innovative Pharmaceutical Firm," in J.A. Cooper, ed., The Economics of Drug Innovation, (Washington, 1970).

28. "An Evaluation of Consumer Protection Legislation," p. 1056.

29. Henry G. Grabowski and John M. Vernon, "Consumer Protection Legislation in Ethical Drugs," American Economic Review 67, no. 1 (February 1977):359-65.

30. Almarin Phillips, "Air Transportation in the United States," Technological Change in Regulated Industries (Washington, D.C.: Brookings, 1971):132-133.

31. Ann F. Friedlander, The Dilemma of Freight Transport Regulation, (Washington, D.C.: Brookings, 1939):97

32. Ibid., pp. 65-99.

33. Ibid., p. 91.

34. Ibid.

35. John W. Kendrick, Postwar Productivity Trends in the United States, 1948-1969 (National Bureau of Economic Research, 1973):94,95; John W. Kendrick, Productivity Trends in the United States, (Princeton: Princeton University Press, 1971):136.

36. Postwar Productivity Trends in the United States, 1948-1969, pp. 94,95.

37. "Air Transportation in the United States," pp. 132-33.

38. W. David Montgomery and James P. Quirk, "Cost Escalation in Nuclear Power," Presented at Conference on Economics of Nuclear Power Regulation, Brookings, Washington, D.C., Feb. 1978.

39. Ibid., p. 25.

40. Ibid., p. 30.

41. Robert Perry, "Development and Commercialization of the Light Water Reactor, 1946-1976", (Santa Monica: RAND Corporation, June 1977).

42. Ibid., p. 29.

43. Ibid., p. 29.

44. Ibid., p. 16.

45. Ibid., p. 10.

46. "Air Transportation in the United States," pp. 132-33.

47. H. Stuart Burness, W. David Montgomery, and James P. Quirk, "The Turnkey Era in Nuclear Power: A Case Study in Risk Sharing Arrangements Involving Regulated Firms," Social Science Working Paper #175, California Institute of Technology, August 1977.

48. Derek J. McKay, "Reliability of Service and Optimal Pricing of Electricity," P-6017 (Santa Monica: RAND Corporation, October 1977); and "Two Essays on the Economics of Electricity Supply," PhD Thesis, California Institute of Technology, August 1977.