THE EFFECTS OF ECONOMIC REGULATION

PAUL L. JOSKOW AND NANCY L. ROSE

Massachusetts Institute of Technology

Contents

1. Introduction 1450
2. The effects of government regulation in general: What are we measuring? 1451
3. Alternative frameworks for evaluating the effects of economic regulation 1453
   3.1. Efficient regulation of natural monopolies 1453
   3.2. "Imperfect" regulation of natural monopolies in the "public interest" 1454
   3.3. Regulation of multi-firm industries 1455
   3.4. The political economy of regulation and its implications 1455
   3.5. Summary 1456
4. Methodologies for measuring the effects of regulation 1457
   4.1. Comparing regulated and unregulated firms and markets 1457
   4.2. Using variations in the intensity of regulation 1460
   4.3. Using controlled environment experiments 1461
   4.4. Structural/simulation models of regulated firms and markets 1462
5. The effects of regulation on prices 1464
   5.1. Franchised monopoly regulation 1465
   5.2. Multi-firm regulation 1468
6. The effects of regulation on static costs of production 1477
   6.1. Franchised monopoly regulation 1477
   6.2. Multi-firm regulation 1480
7. The effects of regulation on innovation and productivity growth 1482
8. The effects of regulation on product quality 1484
9. The distributional effects of regulation and deregulation 1486
   9.1. Profits 1487
   9.2. Factor rent-sharing 1489
   9.3. Transfers among customer groups 1492
   9.4. Transfers among producer groups 1494
10. Conclusions 1495
References 1498
1. Introduction

This chapter discusses alternative approaches to measuring the effects of "economic regulation" and reviews the empirical literature employing these approaches. By "economic regulation" we refer to both direct legislation and administrative regulation of prices and entry into specific industries or markets. We follow conventional treatment in distinguishing economic regulation from a host of other forms of government intervention in markets, including "social regulation" of environmental, health and safety practices, antitrust policy, and tax and tariff policies.

This distinction is at best a practical necessity. Regulatory activities falling in all these categories share common foundations in welfare economics and political economy, and may affect the same economic variables. Firms typically are subject to all these types of intervention, making it difficult to analyze the effects of one type of regulation in isolation from others. However, given the burgeoning literature on regulatory economics during the last fifteen years, an attempt to provide a complete survey of current knowledge on all regulatory effects could easily fill an entire volume. We therefore focus our discussion on effects of economic regulation. This enables us to restrict our attention to an extensive but reasonably well-defined subset of the literature.

Our survey is intended to provide a framework for evaluating and interpreting empirical studies of regulation, a set of guidelines for those embarking on their own empirical investigations, and a review of significant contributions to this literature. The measurement issues that we discuss arise in empirical analysis of all types of government regulation; we therefore have structured the methodological discussion so that it has broad applicability. While the present study is by no means an exhaustive survey of the literature, it includes numerous examples of the use of different types of data and measurement techniques. These are selected to cover a range of industries and time periods sufficient to give the reader a good feeling for what is known, not known, or in dispute. Where particularly useful, we include references to methodological applications from the social regulation literature. Interested readers should consult Joskow and Noll

---

1 Our review reflects the predominantly U.S. focus of the empirical regulation literature: we include references to international or comparative research where relevant. The theoretical and empirical techniques we discuss are broadly applicable to the study of regulation in other countries, though there has been relatively little work of this sort. A partial explanation for this is the tendency for industries that are regulated in the United States to be organized as public enterprises in other countries; see Mitchell and Kleindorfer (1979) and Finsinger (1983). We nevertheless expect major contributions from more extensive analyses of non-U.S. regulatory institutions in the future.
(1981) for a more complete survey of the literature on economic and social regulation through 1980, and refer to the other chapters in this Handbook for analyses of areas beyond the direct focus of this chapter.

The chapter is structured as follows: Section 2 outlines the theoretical bases for identifying and measuring the effects of government regulations. Section 3 develops these frameworks in more detail, discussing theoretical approaches that are of particular relevance for students of price and entry regulation. Section 4 describes four empirical methodologies for measuring the effects of economic regulation. Sections 5 through 9 examine how alternative theoretical frameworks and empirical methodologies have been applied to study the effects of price and entry regulation on each of: prices, costs, technological change, product quality, and the distribution of income and rents. The final section contains a summary and conclusions.

2. The effects of government regulation in general: What are we measuring?

The effects of regulation, whether it is "economic regulation" or "social regulation", are likely to depend on a variety of factors: the motivation for regulation, the nature of regulatory instruments and structure of the regulatory process, the industry's economic characteristics, and the legal and political environment in which regulation takes place. Given the substantial variation in these economic and institutional characteristics, the expected effects of regulation are likely to differ considerably across industries and time. Defining a theoretical framework for analyzing regulation is therefore an important prerequisite to an empirical discussion of regulatory effects. Theory and measurement go hand in hand.

Theoretical research on the economics of government regulation has proceeded from several different perspectives. At one extreme is normative or prescriptive theoretical research, which focuses on when regulation "should" be introduced and what the "optimal" form of regulation is. At the other extreme is a growing body of regulatory research that takes a positive or descriptive perspective, focusing on the economic, political, legal, and bureaucratic forces that lead to government regulation and affect the behavior and performance of regulatory institutions.

Normative research on government regulation can be (roughly) grouped into two branches. The first focuses on identification of "market failures"; that is, imperfections that lead unregulated markets to perform suboptimally relative to some social welfare function (usually the sum of consumer and producer surplus). Natural monopoly, externalities, public goods, information failures, and variations on these themes are standard normative rationales for government intervention into a market economy. The second branch of this literature seeks to develop "optimal" policies for correcting market imperfections (as discussed by Ronald
Braeutigam in Chapter 23 of this Handbook with regard to natural monopolies). Recent research on the incentive properties of different regulatory mechanisms, which explicitly models the information structure of the regulatory environment and the strategic interaction between regulators and those they regulate, has enhanced this literature (discussed by David Baron in Chapter 24 of this Handbook). These extensions recognize that even “good” regulation is imperfect, relative to an ideal in which regulators are costlessly and completely informed about all variables of interest. This helps to set the stage for sound comparative institutional analysis, in which imperfect markets can be compared with imperfect regulation: What is the best that we can do in an imperfect world [Kahn (1979)]?

Positive theories of regulation have matured considerably during the last fifteen years. Historically, positive “public interest” theories of why regulation emerges and how it works were based on normative rationales for optimal intervention [Posner (1974)]; regulators were assumed to maximize social welfare subject to various constraints. In this paradigm, empirical analysis of regulatory effects implicitly becomes both a test of whether or not regulatory institutions are successful in achieving their welfare-maximizing objectives and a basis for quantifying the costs and benefits of regulation. During the past fifteen years, economists have rejected this simplistic model of regulation in favor of richer positive theories of regulatory objectives, processes and outcomes [Stigler (1971), Posner (1974), Peltzman (1976), Wilson (1980), Kalt and Zupan (1984), Noll (1985b)]. These recognize that regulation and regulatory processes respond to complex interactions among interests groups that stand to benefit or lose from various types of government intervention. Specific positive theories of the political economy of regulation then become a possible framework within which the nature and consequences of regulation can be predicted, measured and evaluated. This literature is discussed in more detail in Chapter 22 by Roger Noll in this Handbook.

Empirical analysis of the effects of government regulation can be useful from both normative and positive perspectives. It is, however, important to articulate which framework motivates the analysis. The particular theoretical framework used to develop hypotheses about regulatory effects can have important implications for the nature of the effects one seeks to measure, the formal specification of hypothesis tests, and the collection and use of data. Most importantly, the “effects of regulation” do not mean anything in the abstract. We must ask “the effects of regulation compared to what?”. The theoretical framework that leads to measurement questions generally defines (at least implicitly) what the comparative basis for measurement is. It is essential to specify these underlying assumptions about regulatory and firm behavior as well as the base for comparison. Only from this foundation can one formulate and test precise hypotheses and meaningfully interpret the results.
There are several possible benchmarks against which regulated outcomes can be compared. First, regulatory outcomes may be compared to those that would emerge if the industry performed "optimally", as defined by some welfare criterion. Since these "optimal" outcomes may not in practice be attainable, great caution must be exercised in drawing public policy implications from such comparisons.

Second, regulatory outcomes may be compared to the outcomes that would emerge in the absence of price and entry regulation (deregulation or "no regulation"). Two cautions apply to this choice of benchmark. One should not assume that the unregulated regime would be a perfectly competitive regime; many regulated industries have characteristics that make this assumption quite implausible. Moreover, it is important to define what legal institutions (common law, franchising, etc.) actually exist in the "unregulated" regime. "Unregulated" markets may in practice be markets subject to a different form of regulatory restrictions (e.g. municipal franchise regulation rather than state commission regulation), not markets subject to no regulation at all.²

Third, one set of regulatory institutions may be compared to some alternative set of regulatory institutions. The alternative could involve minor changes within the context of a particular regulatory process – such as introducing more incentives into cost-of-service regulation [Joskow and Schmalensee (1986)] – or more fundamental changes – such as municipal franchise bidding in place of state commission regulation [Demsetz (1968), Williamson (1976)].

If empirical evidence on the effects of regulation is to be useful for normative evaluations of regulation, it is essential that the benchmark used to measure and articulate regulatory effects be clearly defined. Similarly, tests of competing positive theories of regulation rely on measuring the actual effects of regulation, which also requires precise specification of the benchmark against which regulatory effects are being measured. Each of these benchmarks can provide useful empirical evidence for answering normative and positive questions, but only if the benchmarks are articulated clearly.

3. Alternative frameworks for evaluating the effects of economic regulation

3.1. Efficient regulation of natural monopolies

The traditional economic rationale for price and entry regulation is that the production of a particular good or service (or set of goods and services) is

²There is, in reality, no such thing as "no regulation". At the very least firms are subject to common and statutory law institutions affecting property rights, liability, and contracts.
characterized by “natural monopoly” [Schmalensee (1979)]. In this case, a single producer minimizes costs, but an unregulated market would lead to prices or costs that are on average too high and to price structures that may be inefficient. Price and entry regulation may be optimal from a normative perspective if: (a) single firm production of one or more goods minimizes costs, i.e. the production function is subadditive over the relevant output range; (b) a firm with a legal monopoly will choose average prices and profits that are too high (excess profits), and individual prices that may be too high or too low (inefficient rate structure); (c) the threat of entry will not effectively discipline a single supplier; and (perhaps) (d) inefficient entry may occur in the absence of a legal monopoly even if, or because, prices are regulated. This rationale has been used to justify price and entry regulation of electricity supply, natural gas transmission and distribution, telephone service, water and sewer service, and cable television service.3

In these industries, “good regulation” is supposed to: constrain entry so that the economies of single firm production can be achieved; constrain prices so that the firm earns neither excess nor insufficient profits; and regulate the structure of rates so that individual prices are efficient (at least in a second best sense). When we examine the effects of “costless”, well-informed price and entry regulation in industries that are assumed to have natural monopoly characteristics, it seems natural to ask how well regulation achieves these objectives. For example, is the average level of prices constrained below what could be charged by an unregulated monopolist but above the level at which the firm would choose to exit in the long run? Do regulated firms earn normal profits? Is the regulated rate structure efficient in a second-best sense? Empirical analysis of the prices charged and costs incurred by franchised monopolies can, in principle, answer these questions.

3.2. “Imperfect” regulation of natural monopolies in the “public interest”

Regulators are unlikely to be perfectly informed, and regulation is unlikely to be costlessly implemented and enforced. When we expand our normative framework to recognize inherent imperfections, the set of potential regulatory effects becomes quite rich. Analysis of practical, as opposed to ideal, regulation must include explicit consideration of the incentive properties of specific regulatory rules and procedures used to set prices, the dynamics of regulation, the control instruments and information available to regulators, and the responses of regulated firms to all of these. Price regulation that sets rates based on the “cost of service” may distort firms’ input choices [Averch and Johnson (1962), Baumol

3 This line of argument could in principle be generalized to encompass markets characterized by “natural oligopoly” or imperfect competition, although there has been little academic interest in doing so.
and Klevorick (1970), Isaac (1982)], or more generally encourage X-inefficiency [Joskow and Schmalensee (1986)]. Regulation may alter the rate and direction of technological change [Capron (1971)]; distort quality choices [Spence (1975)]; change the financial risk faced by the owners of the firm [Brennan and Schwartz (1982)]; and affect the prices regulated firms pay for inputs [Hendricks (1975, 1977), Ehrenberg (1979)]. Finally, regulation is likely to redistribute income among various interested parties. These distributional effects are of particular interest to those who study the political economy of regulation.

3.3. Regulation of multi-firm industries

Natural monopoly rationales have less inherent plausibility for industries in which several firms, rather than a single franchised monopolist, are allowed to provide service. If the production of some good or service has natural monopoly characteristics, regulatory systems that permit or encourage many firms to provide service, subject to economic regulation, must have some other explanation. Regulation in many such industries has been rationalized by “excessive” or “destructive” competition in an unregulated environment (truckng, banking, airlines) or by “natural oligopoly” (railroads), but these arguments are frequently unpersuasive. Skepticism about the need for regulation based on plausible market imperfection rationales leads naturally to an investigation of the causes of regulation and its effects on prices, profits, and market structure. When competing firms operate in a regulated market, the nature of price regulation itself typically changes, and the variety of possible regulatory effects expands. Regulated prices in industries with multiple competing firms generally are based on some measure of industry average costs rather than the costs of each regulated firm [Daughety (1984)]; non-price competition must be carefully incorporated into the analysis [Joskow and Noll (1981)]. While many of the variables that can be affected by regulation are the same as those described for monopoly markets, the nature of regulatory effects may differ considerably from those that emerge when a single legal monopoly firm serves a particular market.

3.4. The political economy of regulation and its implications

While it may be of interest to compare the effects of economic regulation to the ideal “public interest” regulation, this simplistic “normative theory as positive theory” approach does not provide a sound foundation for positive theories of regulation and its effects. The introduction of price and entry regulation, as well as its structure, operation over time, and effects, reflect a complex interplay among interest groups that stand to gain or to lose from different types of
regulatory intervention—not efforts to maximize the sum of consumer and producer surplus. Regulatory processes and outcomes depend on the magnitude and distribution of the costs and benefits of various regulatory interventions, the structure of the interests groups affected, prevailing economic conditions, and the nature of political, regulatory and legal institutions within which various groups pursue their self-interest. Regulatory outcomes may reflect “public interest” considerations, through the effects of market imperfections on interest group politics, but we cannot assume that this will necessarily be the case.

Viewed from this perspective, the nature and magnitudes of regulatory outcomes can be quite complex. Price and entry regulation may lead to prices that are higher or lower than what would emerge in the absence of such regulations. Rather than seeking to provide consumers with the benefits of economies of scale or scope, regulation may protect firms that are not natural monopolies from the threat of competition and lower prices. Rate structures are likely to reflect interest group politics rather than narrow efficiency criteria. New technologies may be discouraged, rather than encouraged, to protect incumbents. The distributional consequences of regulation and changes in regulation become quite important for understanding the nature of the regulatory process itself and how it changes over time. Empirical analyses of the effects of regulation on prices, costs, income distribution, and the like, become central for distinguishing between competing positive theories.

3.5. Summary

These frameworks suggest a diversity of regulatory effects as well as different motivations for and uses of empirical evidence on the impact of regulation. The specific effects of interest will depend on the theoretical model of regulation and firm behavior that characterizes the industry and regulatory process under study. Different theories of regulation will lead to different predictions about the nature and magnitude of its effects, and the nature of regulatory effects will, in turn, have important implications both for making normative judgments as to whether regulation is “good” or “bad”, and for distinguishing among alternative positive theories of regulation. We will in the remaining sections focus on measuring the effects of economic regulation on the following indicia of firm and/or market behavior and performance:

1) The average price level and the structure of prices (e.g. non-uniform and non-linear tariffs, pricing for multi-product natural monopolies).
2) The static costs of production, including:
   (i) input distortions,
   (ii) X-inefficiency,
   (iii) direct regulatory costs, and
   (iv) input prices paid.
(3) Dynamic efficiency, including the rate and direction of innovation and productivity.
(4) Product quality and variety.
(5) Distribution of income and rents, including:
   (i) profitability of regulated firms,
   (ii) rent-sharing with factors of production,
   (iii) income transfers among customer groups, and
   (iv) income transfers among producer groups.

4. Methodologies for measuring the effects of regulation

There are four basic empirical methodologies for measuring the effects of regulation. Although these approaches are not mutually exclusive, each has particular features that may limit or enhance its value in a specific application. These features are highlighted in our discussion below. We make only limited reference to examples from the literature in this section; a broader discussion of studies employing these methods to measure various regulatory effects is deferred to Sections 5 through 9.

The four approaches we consider are:
(1) Comparing regulated and unregulated firms or markets.
(2) Using variation in the intensity of regulatory constraints.
(3) Controlled environment experiments.
(4) Structural estimation/simulation models of regulated firms or markets.

4.1. Comparing regulated and unregulated firms and markets

A simple approach to measuring the effects of regulation is to compare matched samples of “regulated” and “unregulated” firms (or markets). If the only difference between the samples is the nature of the regulatory constraints the firms are subject to, differences in behavior and performance can be attributed to regulation. This approach may rely either on cross-sectional variation, comparing similar firms operating under different regulatory structures, or on time-series variation, comparing the same firms operating under a changing regulatory environment.

The cross-sectional approach most frequently exploits variation in regulatory environments across states, although other sources of regulatory variation also have been analyzed. Differing regulatory regimes across countries [Moore (1976), Finsinger and Pauly (1986)], exclusion of intrastate firms from federal regulation [Jordan (1970)], and statutory exemptions of some firms or markets within a regulated industry may provide alternative sources of cross-sectional variation. Once variations in regulatory jurisdictions are identified, prices, costs, or other
performance measures are developed; the difference in their levels between firms operating in "regulated" jurisdictions and those operating in "unregulated" jurisdictions is estimated and attributed to regulation. This type of analysis requires both reasonable variation in regulatory regimes and an ability to control for relevant non-regulatory variations across firms. Stigler and Friedland's (1962) seminal paper on regulated electricity prices in "regulated" and "unregulated" states is a classic example of this method of analysis.

The time-series, or "before-and-after", approach exploits variation in regulatory environments over time. This analysis requires identification of a time period (or periods) during which the regulatory regime changes. The behavior and performance of firms or markets before the regulatory innovation is compared to that after the innovation; the difference is interpreted as the effect of regulation. Effects typically are identified from actual responses of performance measures (prices, costs, innovations) to the introduction or elimination of regulation. This requires data prior to and after the change, and ideally would use a fairly lengthy time series to avoid basing conclusions on possible transitional responses. An alternative approach, available for identifying some, but not all, regulatory effects, is to estimate the expected effect of regulatory reforms on performance. This can be accomplished through the use of financial market data and "event study" techniques [see Schwert (1981), Rose (1985a), Binder (1985)].

Either a realization-based or an expectations-based time-series approach is available only when it is possible to identify distinct changes in regulatory regimes and when time-series differences in other relevant variables can be readily controlled for. Peltzman's (1973) study of the effect of FDA regulation of drug efficacy and Rose's (1985a) event study of regulatory rents in the trucking industry are examples of this type of analysis.

Both cross-sectional and time-series analyses involve a common method. First, the dependent variable of interest—such as price, cost, or the rate of technical change—must be defined, and modelled as a function of exogenous economic characteristics that influence performance independent of regulation and a control for the influence of regulation. Regulation generally is measured by a dummy variable indicating whether an observation is drawn from the "regulated" or "unregulated" regime. The effect of regulation is inferred from the sign and magnitude of the coefficient on the regulatory dummy variable.

This dummy variable approach has been used quite widely. Though in theory simple, its implementation and interpretation in practice warrant several cautions. First, it is essential that the differences between the regulatory regimes be carefully articulated. If the "regulated" regime is measured by the existence of state regulatory commissions, for example, it is important that all commissions exercise similar authority over firms' behavior (particularly with respect to price and entry). Treating commissions that have the power to set only minimum (or only maximum) rates as identical to those with authority to set actual rates
introduces noise that may bias downward estimates of the difference between “regulated” and “unregulated” firms. Similarly, if firms operating in states without commission regulation are assigned to the “unregulated” sample, the interpretation of any differences in performance depends critically upon whether firms in these states are completely unregulated or controlled by some other set of restrictions. A clear specification of the alternative regulatory regimes, as well as careful inspection of the institutional structures governing firms’ behavior in each, is critical to this type of empirical analysis.

Second, care should be taken in controlling for non-regulatory differences between firms or markets. The political economy literature, which develops and tests positive theories of regulation, characterizes the introduction, design, and repeal of regulation as an endogenous choice. This suggests a systematic relationship between economic conditions that affect the behavior and performance of firms and the incidence of regulation. Similarly, regulatory changes may follow upheavals in the distribution of costs and benefits, inducing systematic relationships between economic variables and the nature of the regulatory regime in time-series analyses. This argues strongly for developing a detailed model of the interaction of regulatory structures, economic characteristics of firms or markets, and the behavioral or performance measures of interest. This model should then be used to structure empirical tests of regulatory effects. To the extent that there are systematic differences in important economic characteristics of firms and markets between “regulated” and “unregulated” regimes (across jurisdictions or over time), failure to properly control for these differences may bias the measured effects of regulation.

Time-series analyses involve a third complication: determining the date at which regulatory regimes change. Regulatory statutes may directly restrict firm activity or, more typically, may establish a regulatory agency with a broad mandate to develop specific rules, regulations, and procedures. In either case, “grandfather clauses” and implementation or enforcement lags may cause the actual imposition of significant restrictions on activity to lag behind the nominal date of regulation. These difficulties are compounded for deregulation. In this case, substantial revision of regulatory structures may take place through changes in the administering agency’s policies. Moreover, these changes may not occur through formal rulemakings, but may instead be signalled only by decisions in administrative cases. Recent deregulatory experience in airlines, trucking, and banking suggests that such administrative revisions may considerably pre-date congressional legislation.

Expectations-based analyses that use event study techniques must identify the date on which expectations about the regulatory regime change, rather than the date the regime actually changes. The effective date of legislation will be much too late for this type of analysis. In addition, it will be difficult to identify regulatory effects from most congressional activity, as congressional votes tend to
be well anticipated [see the results in Binder (1985)]. Administrative reforms therefore appear more conducive to expectations-based approaches.

These difficulties with dating regime changes suggest the importance of carefully analyzing the sequence of events leading to major reforms in regulation. Reading the statute creating (or abolishing) the regulation under study, reviewing the contemporaneous trade press discussions of the regulation, and examining the administrative rules, policies, and decisions established by the regulatory agency may help to determine a meaningful date for the regulatory change.

Finally, it may be useful to combine the time-series approach with one of the others described below. In particular, the use of panel data on firms or markets can dramatically improve the power of empirical tests of regulatory effects. Rarely do we expect regulation to have the same effects on all firms at all times. By specifying the determinants of differential effects and employing both time-series and cross-sectional variation, we may obtain stronger results. For example, an industry-wide change in regulation may help some firms and hurt others, depending on their particular economic characteristics. Adding cross-sectional data on firms to the time-series analysis and modeling the regulated–unregulated dummy variable as a function of these characteristics could increase the power of statistical tests of regulatory effects [Rose (1985a), Smith, Bradley and Jarrell (1986)]. Similarly, variations in economic conditions or regulatory intensity over time could be used to add a time-series dimension to the cross-sectional dummy variable tests.

4.2. Using variations in the intensity of regulation

In many cases it may not be possible to obtain data on firms or markets that are subject to fundamentally different regulatory regimes. Essentially all states may regulate certain industries, so that distinct cross-sectional variation between “regulated” and “unregulated” environments simply may not exist. There may be no regulatory shock during the time period of interest that makes before-and-after comparisons feasible. In short, we may have observations only on firms and markets subject to qualitatively similar regulatory constraints. This situation clearly is not conducive to the “dummy variable” approach discussed above. Yet there may be quantitative differences in the regulatory constraints applied over time and space that, under particular theories of regulation and its effects, would be expected to yield differences in outcomes in one or more dimensions. These variations may arise from differences in regulatory structures or processes, or from the effects of changing economic conditions on regulation. For example, variations in the “tightness” of the rate-of-return constraint have been used, in the context of the Averch–Johnson model, to predict variations in factor input utilization [e.g. Spann (1974)] and productivity growth [Nelson and Wohar
Ch. 25: Effects of Economic Regulation

Variations in regulatory resources [Norton (1985)], the structure of specific regulatory instruments and procedures (such as fuel adjustment clauses [Gollop and Karlson (1978)], and the treatment of construction work in progress), and independent ratings of the "quality" of regulatory agencies [Navarro (1982)] have been used to examine the effects of regulation on costs and market values. Variations in the nature of environmental restrictions have been used to measure the costs of environmental regulation [Gollop and Roberts (1983)].

Proper application of this approach requires a detailed understanding of variations in regulatory rules and procedures and the specification of a precise model of how these variations affect the behavioral and performance variables of interest. The cautions discussed with regard to the comparative cross-sectional and time-series approaches also apply. The informational requirements for this approach are much stronger than are the requirements for the comparative "dummy variable" approach. Care must be taken to control for differences in economic conditions that may affect measures of regulatory intensity (such as allowed rates-of-return) independently of the regulatory structure.

Interactions of regulation with changing economic conditions may, when properly modelled, provide an additional way of identifying regulatory effects [Joskow (1974), Carron and MacAvoy (1981), Hendricks (1975), Burness, Montgomery and Quirk (1980), Greene and Smiley (1984)]. In particular, certain regulatory constraints may be binding under one set of economic conditions, but not under another. Implementing this approach requires particular attention to the nature of the regulatory process under study and how it works when economic conditions change. Joskow's (1974) model of state public utility commission behavior provides an example of this approach.

4.3. Using controlled environment experiments

Data generated by actual regulatory and economic conditions may not provide sufficient experimental evidence to estimate the effects of regulation. As an alternative to relying on the "natural experiments" provided by actual experience, evidence from controlled experiments is increasingly used to measure regulatory effects [Smith (1982), Plott (1982), Hausman and Wise (1985), Cox and Isaac (1986)]. These experiments are designed to generate data suitable for testing specific hypotheses about the effects of variations in institutional arrangements and public policies. Two types of experimental evidence are potentially available. Field experiments may be designed to study the behavior of real economic agents. In these, economic conditions or institutional structures are varied in

---

4 This is, of course, a potential problem with all econometric work, and not specifically (or more significantly) related to efforts to estimate the effects of regulation.
systematic ways, and behavioral responses are used to quantify the effects of alternative regulatory, public policy, or market arrangements. Field experiments have been conducted to study the effects of a negative income tax, housing subsidy programs, health insurance programs, peak-load pricing [Hausman and Wise (1985)], and the deregulation of the bulk power market [Acton and Besen (1985)]. Field experiments are time-consuming and expensive. Laboratory experiments are an increasingly popular alternative. In these, human (or animal) experimental subjects participate in a set of laboratory “games”, designed to provide the subjects with economic conditions that they would face under various market and institutional arrangements. Institutional details can be varied in a way that carefully controls for other causal variables. This approach is used by Hong and Plott (1982), to examine the effects of regulatory pricing rules on inland barge transportation; Rassenti and Smith (1986), to investigate the performance of unregulated wholesale electricity markets; and Cox and Isaac (1986), to evaluate the effects of incentive mechanisms applied to legal monopolies. While experimental techniques have not yet had a major impact on the study of regulation, this approach is certainly promising.

4.4. Structural / simulation models of regulated firms and markets

In all too many cases, none of the previous approaches can readily be used: there are no significant variations in regulatory regimes, in the intensity of regulatory constraints, or in economic conditions that would enable one to measure directly the effects of regulation on behavior or performance. Controlled experiments may be too expensive or complex to perform. We observe regulatory outcomes, but may not have the sample variation to compare these outcomes to a less regulated benchmark. Even when there is substantial sample variation in regulatory incidence, we may lack confidence in our ability to control for important differences that affect both performance and regulation. In these cases, structural models of behavior or performance, combined with simulation techniques, may provide a means of estimating regulatory effects.

As an example, suppose we are interested in determining whether regulatory agencies constrain the prices that franchised monopolies charge below monopoly levels, what the difference is between regulated and monopoly prices, and whether or not the rate structure is “optimal”. By estimating the demand and cost functions for these firms, we can compare the average regulated price level to the costs of production to determine the relation between prices and costs. We can solve for the monopoly prices under varying assumptions about the degree of price discrimination and entry restrictions, and compare these simulated prices to the actual prices [Smiley and Greene (1983), Greene and Smiley (1984)]. Finally, we can use the system to solve for second-best non-uniform and non-linear
prices, and estimate the welfare gains from more efficient pricing [Brown and Sibley (1986)]. In a similar vein, estimates of firms’ production functions, combined with information on input prices, can be used to test whether regulated firms make cost-minimizing input choices.

The success of this approach depends critically upon the ability to identify and accurately estimate demand and cost functions. This task is in some ways easier for firms operating in regulated industries than for those operating in unregulated industries. Regulatory agencies frequently collect detailed firm-level information on revenue, outputs, input prices and quantities, operating costs, capital stocks, investment, and the like. These data often are available over long time periods, and tend to be comparable across firms and over time due to the agency’s use of a uniform system of accounts. There are, however, a number of potential impediments.

Estimating demand functions for regulated firms or markets should present no unique difficulties. The issues involved in obtaining consistent demand estimates should be independent of regulatory status; the availability of high-quality data should make this task easier to execute in regulated markets. We are not as sanguine about cost or production function estimation. Estimates of production or cost functions from observed combinations of outputs, inputs, input prices, and costs tend to rely on a number of implicit assumptions, including equilibrium conditions and exogenous factor prices. These may be implausible for many regulated markets.

For example, the bulk of utility investments are long-lived sunk investments with putty-clay technology. Once in place, input proportions are close to fixed, implying that input proportions are likely to be unresponsive to changing input prices. Moreover, expected input prices are unlikely to be constant over time. Assuming static input price expectations, or assuming that the firm is in long-run equilibrium with respect to current input prices, as is often done, will yield unreliable results. Similarly, there is considerable evidence that regulation affects input prices [particularly wages – Hendricks (1975, 1977)] and can directly increase costs by restrictions on factor use (such as inefficient route structures imposed on regulated transportation firms). To the extent that one treats factor prices as exogenous, or fails to model explicitly direct regulatory constraints on production decisions, the resulting cost estimates may be quite misleading. This is not to discourage the use of structural estimation/simulation approaches; we find their careful application quite informative. We urge, however, careful consideration of the assumptions implicit in its implementation, and modifications to account for the peculiarities of the particular regulatory process under study where appropriate.

Although this approach generally is quite information-intensive, in some cases very simple calculations can be instructive. For example, under depreciated original cost ratemaking, the relationship between a utility's stock price and
(regulatory accounting) book equity per share varies directly with the relationship between its expected return on investment and cost of capital [see Schmalensee (1986) and the references he cites]. A utility’s price-to-book ratio will exceed (fall below) 1.0 if the firm is expected to earn more (less) than its cost of capital, and will equal 1.0 when the utility is expected to earn exactly its cost of capital. Given certain assumptions about earnings and dividend growth paths, the price-to-book ratio and other financial data can be used to estimate the difference between the expected return on investment and the cost of capital, and inferences can be drawn about whether prices are too high or too low and by how much [see Smiley and Greene (1983) and Greene and Smiley (1984)]. While this approach applies only to regulated firms subject to depreciated original cost ratemaking, analyses based on Tobin’s “q” could provide similar inferences independent of the form of regulatory ratemaking [Lindenberg and Ross (1981), Salinger (1984), Smirlock, Gilligan and Marshall (1984), and Rose (1985b)].

Another simple application of structural models uses asset pricing theory. Regulation may create assets that have value only in a regulated environment, such as operating certificates for regulated trucking companies (i.e. licenses to operate in the specified market), taxicab medallions, radio and television broadcast licenses, crude oil entitlements, and state liquor licenses [Schwert (1981)]. If these assets are traded, their prices will reflect the capitalized value of expected regulatory rents accruing to the holder. Measuring asset values becomes complicated if their sale is bundled with other assets (as is the case with broadcast licenses and taxicab medallions in many jurisdictions); interpreting their value is difficult if they reflect an allocation of scarce resources (such as the broadcast spectrum) as well as regulation-imposed scarcity. These and other issues are discussed at length by Schwert (1981). Despite potential complications, regulatory assets permit a fairly clean test of profitability effects.

5. The effects of regulation on prices

There has been extensive empirical research on the effects of economic regulation on the average level and structure of prices; no simple generalization emerges from this work. Depending on the industry, type of regulation, time period, and norm for comparison, regulation has been shown to increase prices, decrease prices, distort the structure of prices in a variety of different ways, and sometimes to have no significant effect on prices at all. The implications of regulation for

5A price-to-book ratio greater than one does not necessarily imply that regulation is too lax. The combination of a modest wedge between prices and costs with regulatory lag may promote static and dynamic cost minimization.
prices therefore depend on the regulatory and economic characteristics of the particular industry being studied.

5.1. Franchised monopoly regulation

We first address research that focuses on "natural monopoly" industries; those for which price regulation combined with de facto franchise exclusivity has been justified on natural monopoly grounds. These include electricity, natural gas distribution and (perhaps) transmission, telephone service, and water and sewer service.\(^6\) Despite the central role of "natural monopoly" in normative theories of regulation, surprisingly few studies have estimated the effects of regulation on the level and structure of prices charged by franchised monopolies. Existing work on price level effects has focused on electricity prices, while analyses of rate structures have covered both electricity and telephone pricing.

Stigler and Friedland (1962) provide the first systematic econometric study of the effects of state commission regulation on electricity prices. They use a comparative cross-sectional methodology to measure average electricity prices in states with state commission regulation of electricity rates relative to prices in states without such regulation, controlling for differences in production costs. The results indicate small and generally insignificant negative effects of regulation on prices, and may suggest a slight increase in the constraining effect of regulation over time. The interpretation of these results highlights two methodological issues. First, insignificant results do not imply that state-regulated prices were identical to unconstrained monopoly prices. During the time period studied by Stigler and Friedland, state commission regulation typically replaced municipal franchise regulation and established clear compensation rules in the case of municipal takeover. The "regulated" dummy variable therefore measures the difference between state regulation and municipal regulation or ownership, not "no regulation". There is little reason to expect this difference to be significantly positive.\(^7\)

\(^6\) These services are not always provided by private for-profit firms, particularly outside the United States. In Europe, government-owned enterprises dominate these industries. In the United States, water and sewer service is typically, but not always, provided by government agencies; municipal and cooperative distribution companies account for about 20 percent of electricity sales [Joskow and Schmalensee (1983)]; gas distribution service is sometimes provided by municipal utilities; and local telephone service is sometimes provided by cooperatives.

\(^7\) The fact that Samuel Insull, the leading electric utility entrepreneur of the day, was a leading proponent of state commission regulation [McDonald (1962)] suggests that state regulation may have led to higher electricity profits. Whether this resulted from higher prices to consumers or lower payments to regulators (in the form of non-price concessions or bribes) is unclear.
Second, caution should be exercised in generalizing these findings beyond the time period they cover. In particular, the demand faced by a franchised electric utility in the 1920s and 1930s was probably much more elastic than it was in later years, and the unconstrained monopoly price much lower relative to the cost of production.\(^8\) Alignment of “regulated” and “unregulated” prices may reflect low monopoly power as much as ineffective regulation.

Application of this approach to study contemporary effects of public utility regulation is essentially precluded by the pervasiveness of state commission regulation.\(^9\) One natural alternative to the comparative cross-sectional approach is the structural/simulation methodology, which uses estimates of the cost and demand functions for public utility service to calculate prices under assumptions about industry structure, behavior and performance in the absence (or with a different form) of price and entry regulation. Numerous studies have estimated demand functions for electricity at different levels of aggregation [Taylor (1975), Baughman, Joskow and Kamat (1979)]; others have estimated electric utility cost functions [Christensen and Greene (1976) and the references they cite]. Contemporary estimates of the long-run demand elasticity for electricity average about unity; the short-run demand elasticity is much smaller. This suggests that prices could be profitably raised (to equate MR and MC), implying that regulation constrains electricity prices below monopoly levels.

Demand and cost information can be used to compute unconstrained monopoly prices, as well as first-best and second-best efficient prices, which can then be compared to actual regulated prices. Using this type of approach, Smiley and Greene (1983) and Greene and Smiley (1984) find that unconstrained monopoly prices for electricity are 20–50 percent higher than actual regulated prices. Baron and Taggart (1977) introduce an explicit regulatory constraint into a model of electricity production cost characteristics and electricity demand. Using firm-level data for 1970, they also find that regulation constrains electricity prices below the pure monopoly level. In contrast, Breyer and MacAvoy’s (1974) application of this approach to natural gas pipelines suggests little, if any, effect of pipeline regulation on prices.

Utilities’ market-to-book ratios for common equity may also provide information on regulated prices, as described in Subsection 4.4. Market-to-book ratios have varied tremendously over time and space [Greene and Smiley (1984)].

\(^8\)Electricity use prior to the 1930s was much more discretionary than it is today. Real electricity rates were quite high; many customers, especially outside of urban areas, had no electricity service; and residential electricity use was largely restricted to lighting. Although industrial use was rapidly expanding, as late as 1925 more than half of industrial electricity consumption was accounted for by self-generation [Edison Electric Institute (1974)].

\(^9\)Some studies have tried to exploit the remaining variation in the interstate incidence of commission regulation to estimate the effects of regulation on the costs of production [Petersen (1975)] and on systematic risk [Norton (1985)]. The idiosyncrasies associated with the few states not adopting commission regulation by the 1960s make these results difficult to interpret.
Electric utility market-to-book ratios were generally far above unity in the 1950s and 1960s, fell below unity by the late 1970s, and presently are slightly above unity. There is substantial systematic variation in market-to-book ratios across utilities, at any point in time, which may be a function of firm-specific economic characteristics such as construction program size, magnitude of nominal rate increases, nuclear plant under construction, and excess capacity. These suggest substantial variances in regulatory price effects, although there have been few efforts to relate these variations to regulatory and economic conditions [see Greene and Smiley (1984)]. The endogeneity of allowed rates of return has been empirically modelled by Joskow (1972) and Hagerman and Ratchford (1978). Joskow (1972) finds that allowed rate of return decisions reflect firm financial performance and economic conditions. Hagerman and Ratchford (1978) examine the effects of both financial and political variables on agency decisions, and conclude that economic variables are of most importance in determining allowed returns. These studies highlight the time-specific character of studies of regulatory price effects: results are likely to depend critically upon the economic conditions over the sample period.

Joskow’s (1974) study develops an explicit model of the link between economic conditions and regulatory price effects. His model of the behavior of state public utility commissions predicts that the nature of regulatory constraints will vary directly with prevailing economic conditions. With constant or declining nominal costs, the model predicts that regulation will be essentially non-binding. As nominal costs rise during inflationary periods, regulators attempt to minimize or delay price increases, and regulation becomes increasingly constraining. Joskow uses the predictions generated by this model to test the effects of state public utility regulation on electricity prices and electric utility financial performance during different economic regimes. Regulation seems to bind most when nominal costs are rising quickly. These results are confirmed by Greene and Smiley (1984) using more recent data.

Experimental techniques have been used in a limited way to learn more about the price effects and desirability of regulation compared to an unregulated regime. The Federal Energy Regulatory Commission recently sponsored a deregulation experiment for certain short-term wholesale electricity transactions in the southwestern United States [Acton and Besen (1985)]. The experiment did not include a control group, but relied instead on a before-and-after comparative approach. Removing regulatory restrictions appeared to have little effect on prices and quantities, although this may be a consequence of the minimal regulation currently imposed on wholesale transactions of the type covered by the experiment [Joskow and Schmalensee (1983)].

Theoretical research on efficient pricing for natural monopoly services (peak-load pricing, Ramsey pricing, non-uniform pricing, etc.) has led to considerable empirical interest in the rate structures established by regulators. This work
focuses on both the practical implementation of efficient pricing schemes [Joskow (1976), Turvey (1968), Nelson (1964)] and evaluations of how closely regulated prices conform to these ideals.

The Department of Energy sponsored numerous field experiments during the 1970s, with varying degrees of success [Aigner (1985)]. In one of the better analyses of the experimental data, Acton and Mitchell (1980) use evidence from a Los Angeles experiment to estimate the welfare gains associated with peak-load pricing of electricity for residential customers. They find that, after accounting for the additional costs of metering, the welfare gains from time-of-day pricing are relatively small and are limited to consumers using relatively large amounts of electricity.

Experimental techniques have been used extensively to analyze electric utility rate structures. Brown and Sibley (1986) employ structural estimates of demand and cost to infer the price changes and welfare consequences associated with efficient pricing of telephone service. Mitchell, Manning and Acton (1978) use international variations in the use of peak-load pricing for electricity to identify industry responses to time-of-day pricing. They rely on differences between regulatory outcomes in the United States, where peak-load pricing was rarely used prior to the late 1970s [Joskow (1979)], and public enterprise outcomes in Europe, where extensive use was made of peak-load pricing after World War II.

5.2. Multi-firm regulation

Considerably more research has been devoted to estimating the price effects of economic regulation in multi-firm industries such as airlines, trucking, railroads, property/liability insurance, hospitals, natural gas and petroleum wellhead production, certain agricultural commodities, and professions subject to state licensing restrictions. Many of these industries have undergone substantial regulatory reform or deregulation since the mid-1970s. This provides a series of natural experiments for studying the effects of regulation using a time-series approach; the deregulation “shock” makes it possible to observe changes that take place when regulatory constraints on prices and entry are removed or changed in fundamental ways.

In contrast to public utility regulation, regulated prices in many of these industries – particularly airlines, trucking, railroads, natural gas, and property/liability insurance – were in principle based on average cost characteristics for groups of firms, rather than the costs of individual firms. In others, such as licensed professions, entry or supply is regulated and prices are determined by market-clearing conditions. In still others, regulators fix price, but leave supply essentially unrestricted, to be determined by market conditions. Under all these structures, regulation may not entirely eliminate competition, but may instead
channel it in directions other than prices. As a result, the effects of regulation on prices may be intertwined with its effects on costs and product quality. We will, however, attempt to focus on price effects in this section, and defer most discussion of cost and quality effects to subsequent sections.

5.2.1. Airlines

In the airline industry, the Civil Aeronautics Board (CAB) set prices and restricted entry into new markets, but firms could still compete in non-price dimensions, particularly on service quality. Douglas and Miller's (1975) analysis of the effects of CAB regulation is particularly noteworthy. This study develops a useful theoretical model of regulation and applies several of the empirical approaches described in Section 4 to measure its predicted effects. Douglas and Miller's work, along with related studies by Levine (1965), Jordan (1970), Eads (1975), and Keeler (1972), shows that, on average, both regulated rates and airlines' service quality choices were too high [see also Bailey, Graham and Kaplan (1985, ch. 1)]. Douglas and Miller compare two aspects of service quality—the proximity of flights to passengers' desired departure times and the probability of being able to obtain a seat on short notice to estimates of the marginal valuation of service quality. Their results indicate that airlines provided higher quality (more costly) service than consumers desired at the margin, implying that average prices were too high. Douglas and Miller also use a "regulated–unregulated" comparative approach [relying on Jordan (1970)], to compare regulated interstate fares with unregulated intrastate fares in California. They report that for comparable routes, CAB regulated fares were higher and average load factors were lower than were unregulated California fares and load factors.

Studies of airline regulation also find that fare structures deviate from efficient pricing rules. Because the CAB's fare formula was not sufficiently sensitive to the effects of market density and distance on costs, fares on longer and denser routes were too high relative to costs, and those on shorter and less densely traveled routes were too low. Regulators also discouraged or prohibited the use of peak-load prices and non-uniform rates, even where these could be justified by cost and demand conditions [Bailey, Graham and Kaplan (1985)]. The deregulation of the domestic airline industry in the late 1970s provides an excellent natural experiment for testing hypotheses about the effects of regulation. While evidence on the effects of deregulation is still being accumulated and analyzed, several recent studies shed some light on the effects of airline regulation on prices.

---

10 These quality measures are used directly in their simulation analysis, although their other work uses average load factors as a proxy. Average load factors are used to measure quality in most other pre-deregulation studies of airline regulation.

Bailey, Graham and Kaplan compare observed fares to what fares would have been had the CAB fare formula continued to be used. They also compare price changes and cost changes. They find that, on average, fares increased less than did average operating costs and less than they would have under regulation [Bailey, Graham and Kaplan (1985, p. 61)]. Similar results are found by Call and Keeler (1985). Morrison and Winston (1986, pp. 22–24) compare actual 1977 fares to predictions of what 1977 fares would have been under deregulation, and conclude that deregulated coach fares would be on average 10 percent higher, while average discount fares would be 15 percent lower under deregulation.

The permanence of these price declines remains questionable, however. Prices tend to be lower relative to costs in markets that are less concentrated and in markets that are served by one or more of the low cost “no-frills” entrants [Graham, Kaplan and Sibley (1983), Call and Keeler (1985)]. Call and Keeler suggest that the post-deregulation competitive environment is characterized by strategic oligopoly behavior, not contestability. Today, many of the new entrants are no longer independent players in the market. These concentration and entry results are inconsistent with the view that the airline industry is “contestable” and suggest that the recent wave of airline merger activity might reverse the early price declines observed under deregulation.

Deregulation also increased the variance of prices across markets. Fares on long-haul routes and denser routes fell considerably relative to the fares simulated by the CAB’s ratemaking formula, while fares on short-haul and less dense routes increased above the levels that would have prevailed under regulation [Bailey, Graham and Kaplan (1985, pp. 54–56)]. Routes that rely heavily on tourists, who arguably have much higher demand elasticities than do business travellers and are more flexible in their choice of departure times, exhibited especially low fares. Whether this reflects efficient peak-load pricing or price discrimination has yet to be determined.

5.2.2. Surface freight transportation

Studies of regulatory effects on prices in the surface freight transportation sector have applied a number of methodologies. Differences in regulatory structures between railroads and trucking appear to have motivated choices of different theoretical frameworks. Trucking regulation was characterized by restrictive entry policy and collectively set rates with rigid price floors; empirical research has focused on price and profit effects, and testing whether regulation cartelized the industry. Rail regulation has been associated with restrictive exit, merger, and maximum rate policies. Empirical rail research has focused on possible modal
choice distortions and welfare effects of the regulated rate structure relative to optimal regulation.

The comparative approach has been used extensively in trucking studies. Snitzler and Byrne (1958, 1959) provide one of the earliest regulatory applications of the comparative time-series approach in their studies of the effect of regulation on trucking rates for certain agricultural products. They find that rates for a variety of food products fell by an average of 19–36 percent when a series of court decisions exempted their shipment from price and entry regulation. Sloss (1970) used inter-provincial differences in Canadian trucking regulation with a comparative cross-sectional approach to measure rate effects. He found that average revenue was roughly 7 percent lower in “unregulated” provinces, although the limited cross-provincial variation in regulation and the potential correlation of economic environments with regulation create some difficulties in interpreting this result. Moore (1976) found larger rate differences in his comparative study of trucking regulation in the United States and Europe, but the absence of controls for differences in economic environments may confound the results.

A second set of studies uses asset market data to measure regulatory price effects. The operating certificates required by the ICC to serve a particular market are a classic example of a regulatory asset: they have no intrinsic value apart from the value of regulatory price and entry restrictions. Therefore, significant positive values reflect regulatory rents, implying supra-competitive pricing. Breen (1977), Moore (1978), and Frew (1981) all find evidence of substantial certificate values in the trucking industry. While a number of complications limit confidence in any particular point estimate of the aggregate value of operating certificates, the results of these analyses provide strong evidence of regulatory increases in trucking rates.11

Trucking regulatory reforms during the late 1970s and early 1980s provide time-series variation that could be used to identify price effects. Unfortunately, there have been few systematic econometric studies of price behavior over time. Blair, Kaserman and McClave’s (1986) comparative time-series study of Florida trucking rates suggests that intrastate deregulation reduced rates by roughly 14 percent. Moore (1986) reports some evidence of price declines coincident with deregulation, but does not attempt to control for changing economic conditions (particularly the 1981–82 recession). Rose (1985a) provides indirect evidence on regulatory price effects, using the variation in regulatory regimes with an event study methodology to test models of regulation. Her evidence is consistent with

11Only a small percentage of certificates ever trade; purchases of certificates may be tied to purchases of other firm assets; prices may reflect the value to the purchaser of improving his network configuration or system profitability, not only the excess profits on the certificated route; and certificate sales typically are bundled, with resulting prices reflecting a mixture of routes and commodities with different characteristics.
the "cartelization" view of trucking regulation, and is suggestive of deregulation-induced price declines. Consistent with earlier studies, the effects appear to be largest in the less-than-truckload sector.

Evidence on price effects of ICC rail regulation typically has been ancillary to measurement of the efficiency losses associated with modal choice and output distortions of ICC regulation. The dominant methodology combines structural estimation of cost and demand conditions with simulation techniques. These analyses highlight the importance of clearly specifying the framework within which regulatory effects are to be measured; in particular what is the alternative to current regulation. Numerous studies, because of their focus on efficient allocation of traffic, compare regulated prices to marginal cost prices [Boyer (1977), Levin (1978, 1981), Friedlaender and Spady (1981), Winston (1981), Braeutigam and Noll (1984)]. This benchmark is appropriate if one is interested in understanding deviations from first-best outcomes; however, it may not be feasible, nor is it indicative of unregulated rates. If pricing at long-run marginal cost results in losses [Friedlaender and Spady (1981), Keeler (1983)], first-best outcomes are unlikely to be attainable. This suggests second-best (Ramsey) prices as an alternative benchmark [Winston (1981)]. Since railroads are likely to possess market power in at least some markets, unregulated prices may deviate considerably from marginal cost; we therefore may consider unregulated prices as a third benchmark [Levin (1981)]. The choice depends critically upon what questions we wish to answer.

Most of these studies estimate modal choice demand functions for commodity groups, then either estimate rail and trucking cost functions (Friedlaender and Spady, Winston) or use ICC cost data [Boyer (1977), Levin (1978, 1981)]. These estimates are used to simulate rates and traffic divisions under various behavioral assumptions. Studies that compare regulated rail rates to marginal cost prices typically find that average regulated rail rates are above marginal costs [Boyer (1977), Friedlaender and Spady (1981), Levin (1978, 1981), Winston (1981), Braeutigam and Noll (1984), Keeler (1983)], although there is substantial variation across commodities. These studies find large welfare losses from existing prices relative to first-best prices: estimates center in the range of $900 million to $1.8 billion annually in 1986 dollars. [Braeutigam and Noll's (1984) critique of the methods employed in some of these studies suggests that true welfare losses may be even higher.]

Both regulated and marginal cost rates result in substantial losses for railroads. Levin (1981), for example, estimates railroad rates of return on book value or replacement cost of assets at 0.75–1.6 percent under marginal cost pricing and at roughly 2 percent under ICC regulated rates. This suggests that regulation has held average rates substantially below unregulated levels. Levin's (1981) simulation of unregulated rail rates under a variety of assumptions about rail competitiveness and regulatory cost effects confirms this. For most plausible scenarios,
average rail rates would increase under deregulation, with the extent of increase most dependent on the degree of interrailroad competition. The results vary substantially across commodities, suggesting considerable regulatory distortions of the rate structure. Boyer (1981) also analyzes rate structures, by relating regulated rail and trucking rates to characteristics of shippers and shipments. He finds that many cost-based characteristics do not influence rail rates, and argues that the pattern of rates suggests a model of "equalizing discrimination", an ICC policy of equalizing conditions between "advantaged" and "disadvantaged" shippers. His analysis of trucking rates reveals patterns consistent with a cartel model of regulation.

The rail industry, like trucking, underwent substantial regulatory reform during the late 1970s and early 1980s. MacDonald (1986) uses this time-series variation to identify regulatory effects on grain transportation. He notes the difficulty of controlling for other changes taking place during this period (such as declining export demand), but argues that grain shipment rates appear to decline during the 1980s even after allowing for these effects. This is broadly consistent with the results of the earlier simulation-based studies. Friedlaender (1988) also provides an analysis of rail rates under deregulation. Further research along these lines seems desirable.

A final transportation mode — inland barge transportation — has attracted relatively little recent interest, although it has been the subject of one of the few experimental studies directly related to regulatory issues. Hong and Plott (1982) use experimental techniques to compare the properties of negotiated prices to those of a posted price system. This experiment was intended to advise the ICC in their consideration of a regulation that would require carriers on inland waterways to file proposed rate changes with the ICC at least fifteen days before they take effect. Hong and Plott's results suggest that a pre-notification policy leads to higher rates, lower volumes, and less efficiency than a policy that allows carriers to file and use new rates immediately.

5.2.3. Insurance

The effects of rate regulation on property and liability insurance premiums for personal lines (auto, residential fire, homeowners' insurance) have been studied extensively since Joskow's (1973) paper on regulation and competition in the property and liability insurance industry; see Ippolito (1979), Samprone (1979), Smallwood (1975), Walter (1979), Williams and Whitman (1973), Kunreuther, Kleindorfer and Pauly (1983), U.S. Department of Justice (1977), and see Harrington (1984) for a recent survey.\(^{12}\) Much of this literature uses a comparative approach, exploiting either cross-sectional differences in regulation across

\(^{12}\) There has been almost no analysis of the effects of regulation on commercial lines of insurance.
states (many states have introduced open competition laws) or time-series variation (much of the deregulation of rates took place after the mid-1970s). The studies show that the effects of rate regulation have varied widely over time and space.

Joskow’s (1973) study concluded that the provision of most lines of insurance was structurally competitive, with many suppliers, easy entry and low concentration. Prior approval price regulation appeared to be a “producer protection” initiative by the insurance industry, undertaken after a 1944 Supreme Court decision ruled that the antitrust laws applied to the joint ratemaking activities of the insurance industry. (Earlier court decisions had concluded that insurance was not covered by the antitrust laws.) Joskow examined the effects of New York’s introduction of an open rating system in the early 1970s, and found that it led many firms to set rates different from the “standard rates” normally filed by insurance rating bureaus and approved by regulators. He also suggested that price competition from lower cost direct writers [Cummins and Vanderhei (1979)] was partially restricted by regulation, leading to higher rates (and costs) on average. Finally, he hypothesized that non-price competition and excess demand for insurance could be a consequence of regulatory ratemaking procedures that fixed prices but not competition in other dimensions.

Harrington (1984) surveys the voluminous literature since Joskow’s 1973 study. The findings of these later studies are mixed. Some discover that rates are lower in unregulated states than in regulated states; others find no effect of regulation; still others find higher rates in unregulated states. Rates charged by direct writers are almost always lower than rates charged by insurers who distribute insurance through agents, and direct writers appear to increase their market shares when pricing constraints are removed. Stringent rate regulation also has adverse supply side effects, forcing some consumers into residual markets (e.g. assigned risk pools) for insurance.

This variation is not terribly surprising, given dramatic changes in the economic conditions faced by property/liability insurance firms. It may be that during the early 1970s, before the acceleration of inflation, regulated rates were higher than competitive market levels and deregulation led to rate reductions. By the late 1970s and early 1980s, rapid inflation could easily have led to a situation in which “regulatory lag” prevented regulated rates from keeping pace with increasing costs, depressing regulated prices relative to competitive levels. This variation would be consistent with the interaction between economic conditions, regulation, and regulatory lag described by Joskow (1974) in the context of electric utility regulation.

---

13Regulatory lag arises when regulators do not continuously adjust prices as cost and profits change. Some regulatory lag is a natural outcome of the administrative process; it can also reflect strategic and political motivations. Rapidly changing nominal costs increase its effects on price/cost relationships. Regulatory lag may have important effects on the incentives firms have to minimize costs and on the quality of service [Joskow and Schmalensee (1986)].
The variation in results may also arise from the inadequacy of a simple dummy variable for capturing differences in regulation across states and the difficulty of measuring prices accurately. In the insurance market, the intensity of regulation varies even among "regulated" or "unregulated" states. Some insurance commissioners have been very consumer-oriented; others have focused on protecting producers. A dummy variable cannot capture these variations. Furthermore, most of these studies approximate prices by loss ratios instead of using actual prices. With free entry and exit, however, long-run loss ratios will tend to equalize, independent of regulatory intensity. Regulatory effects will be manifested by variations in the range of available insurance products, excess demand, and differences in the quality of service. The difficulty of measuring and controlling for these non-price dimensions of insurance output contributes substantially to the wide range of results that have been obtained.

5.2.4. Energy

The effects of administrative regulation of field prices of natural gas have been studied extensively. Studies have examined the effects of area rate price ceiling regulation beginning in the early 1960s [MacAvoy (1962, 1971), Breyer and MacAvoy (1974), MacAvoy and Pindyck (1973), Pindyck (1974), Brown (1970), Erickson and Spann (1971)] and the effects of regulations introduced in 1978 by the Natural Gas Policy Act (NGPA) of 1978 [Broadman and Montgomery (1983), Braeutigam (1981), Braeutigam and Hubbard (1986), Kalt and Leone (1986)]. Most of this analysis uses a structural approach, drawing on econometric models of demand and econometric or engineering models of supply to compare regulated outcomes with simulated market outcomes.

There is reasonably broad agreement that the Federal Power Commission's area ratemaking approach, introduced in the early 1960s, kept field prices too low to clear the market, resulting in shortages and inefficient utilization of natural gas. The NGPA tried to correct some of the resulting distortions by instituting an incredibly complex system of field price regulations. These included: extending regulation to intrastate gas; maintaining a uniform national ceiling price at a level far below market-clearing levels; raising ceiling prices for certain supplies of "new gas"; deregulating certain categories of high cost gas; indexing ceiling prices; and phasing out price regulation in 1985 and 1987 for selected categories of gas [Braeutigam and Hubbard (1986, table 4)]. The immediate effect of the NGPA, exacerbated by the dramatic oil price increase in 1979 and 1980, was a sharp increase in prices for unregulated categories of gas. There also is evidence that price constraints led to increased non-price competition in the form of longer contracts and larger take-or-pay provisions for price controlled gas [Hubbard and Weiner (1984), Masten and Crocker (1985), Crocker and Masten (1988)]. During the early 1980s, there was substantial concern that the scheduled 1985 deregulation of certain categories of gas provided for by the NGPA would
lead to a sudden increase in prices for these categories of gas, since ceiling prices were far below the Btu equivalent price for oil [Braeutigam (1981, p. 180)]. Instead, the unexpected collapse of oil prices in 1985 and 1986 left many pipelines and distributors with high-cost contracts for gas that they could not market in competition with oil.

5.2.5. Other industries

In a number of industries, supply or entry is directly regulated, but prices are determined by the interplay of demand and (constrained) supply. A prominent example of this type of regulation is agricultural marketing orders. Supply side regulations in agriculture, especially in milk marketing, have been the subject of considerable research [see MacAvoy (1977), Masson and Debrock (1980, 1982), Ippolito and Masson (1978), Shepard (1986)]. These studies tend to rely on structural estimation and simulation approaches. They find that marketing orders, which permit producers to market all output through a common agency, restrict the supply of milk or other agricultural commodities available in certain “prime” markets (e.g. class A fluid milk). This tends to raise prices in primary markets and shift substantial supplies into secondary markets (e.g. powdered milk), thereby forcing secondary market prices below competitive levels. The efficiency losses from these regulations can be quite large (Ippolito and Masson, MacAvoy).

Licensed occupations are a second sector in which regulation imposes entry and related supply-side restrictions without directly regulating price. Interest in professionals subject to state licensing requirements was stimulated by Benham’s (1972) study of eyeglass prices, which found that prices were lower in states that allowed professionals to advertise than in states that restricted advertising. Kwoka (1984) finds similar effects in an analysis of advertising’s impact on optometrists’ prices and qualities. Cady (1976) found a similar result for prescription drugs.

State licensing laws by their very nature provide at least some restriction on entry. However, the strength of these restrictions varies across states, making it possible to measure regulatory effects from interstate variations in regulatory intensity. Shepard (1978) uses an interstate comparative approach to measure dentists’ fees in states with licensing reciprocity (i.e. states that waive licensing requirements for dentists licensed in other states) relative to fees in states without reciprocity. He found that states that did not provide for reciprocal licensing (implying more restrictive entry constraints) had dental fees 12–15 percent higher than those in states with reciprocity. Conrad and Sheldon (1982) expanded this analysis to consider commercial practice restrictions on advertising, the number of offices a dentist could operate, and the number of hygienists a dentist could employ, as well as reciprocity regulations. They found that reciprocity restrictions increased fees, but did not find any systematic effects from the other types of

6. The effects of regulation on static costs of production

When regulation is less than ideal, as it necessarily is in practice, its implementation may give rise to a host of production distortions. In this section, we analyze regulatory effects on production costs, focusing on static production efficiency issues: cost-minimizing input proportions, X-inefficiency, and direct regulatory cost increases. The effects of regulation on dynamic efficiency (productivity growth and technical change) are covered in Section 7. We begin with a discussion of the evidence on franchised monopoly industries and then consider studies of multi-firm industries.

6.1. Franchised monopoly regulation

Most empirical research on production cost effects of franchised monopoly regulation focuses on the electric power industry. The Averch–Johnson (A–J) model [Averch and Johnson (1962), Baumol and Klevorick (1970), Bailey (1973)], in particular, has motivated a substantial amount of empirical analysis. This simple model of rate-of-return regulation yields clear empirical predictions regarding input utilization: overcapitalization when the allowed rate-of-return exceeds the utility’s cost of capital. Papers by Spann (1974), Courville (1974), Petersen (1975), Atkinson and Halvorsen (1980), and Nelson and Wohar (1983), among others, have all sought to determine whether electric utilities employ inputs efficiently and whether regulation induces systematic biases in input mix. Baron and Taggart (1977) test similar regulatory effects in a model that differs from the A–J model and more carefully accounts for financial and tax considerations. Rothwell (1985) provides indirect evidence on A–J effects when he finds that utilities’ technology choices are most consistent with a model of net present value (profit) maximizing behavior. These papers typically employ cross-sectional firm- or plant-level data (sometimes both and sometimes as a panel data set) and use variations in regulatory intensity to identify potential effects on production decisions. Nelson and Wohar use aggregate time-series data for the electric power industry, and Petersen also uses a comparative cross-sectional approach, estimating costs in “regulated” versus “unregulated” states.

Spann (1974), Courville (1974), and Petersen (1975) find evidence of significant overcapitalization, consistent with the predictions of the A–J model. Nelson and Wohar’s (1983) results are unstable and raise questions about the model and/or the data they utilize. For most time periods their estimates imply a value for the
regulatory constraint that is negative, despite the fact that the model they use requires that the value lie between zero and one. For the 1974–78 period they find a significant A–J effect, but given the implausible results for the other periods, the reliability of their specifications and data is uncertain. Baron and Taggart (1977) find undercapitalization, which is inconsistent with the A–J model. Atkinson and Halvorsen (1980) find overcapitalization with regard to the capital/labor ratio, inefficient input utilization with regard to the fuel/labor mix, but efficient input utilization with regard to the capital/fuel ratio. Since capital and fuel account for the bulk of electricity production costs, this suggests that input inefficiency due to A–J type regulatory biases is unlikely to be large. Atkinson and Halvorsen attribute this “negative” result for capital/fuel and fuel/labor to the use of automatic fuel adjustment clauses, and suggest that countervailing distortions are at work. This rationalization seems somewhat implausible, particularly in light of empirical studies of fuel adjustment mechanisms such as Gollop and Karlson (1978), discussed below.

These studies are subject to a number of potential weaknesses. First, the A–J model and other theories of regulation and its effects are theories of the firm, not theories of the plant. One must be careful using plant-level data to test firm-level theories. Plants are not built and utilized in isolation, particularly in an electric power system. Investment and utilization decisions at the plant level should be evaluated in the context of the overall optimization problem for the firm. Second, we have serious reservations about basing production efficiency conclusions on what are essentially ex post cost functions, particularly when long-lived sunk investments are important, input prices are uncertain and change over time, technological change in generating technology is taking place [Joskow and Rose (1985a, 1985b), and plant efficiency varies over time [Joskow and Schmalensee (1985)]. Investment decisions should be evaluated by the expected present discounted value of available alternatives at the time the investments are made, not after the fact. The proper way to evaluate fuel and labor utilization decisions ex post is to take the capital stock as given. Except for Baron and Taggart, the literature exhibits little sensitivity to these considerations.

Finally, contemporary analyses of electric utility regulation that rely on comparisons between regulated and unregulated states may have serious identification problems. Few states had not introduced commission regulation by 1970, and these may have atypical characteristics that make it difficult to identify specific regulatory effects. Petersen’s comparison of costs between regulated and unregulated states, for example, is quite sensitive to the fact that seven of his nine “unregulated” plants are gas-burning plants located in Texas [McKay (1976)]. These had very low costs during the 1970s, reflecting both locational advantages and the availability of intrastate gas. Attributing to the presence or absence of regulation the cost differences between these plants and coal or oil plants located
Ch. 25: Effects of Economic Regulation

elsewhere is implausible. We are inclined to agree with McKay’s (1976) conclusion that the results of these studies are unreliable.

Despite increasing theoretical interest in incentive effects of regulation [see, for example, Baron and Besanko (1984), Shleifer (1985), Laflont and Tirole (1986) and the studies cited in Chapter 24 by Baron in this Handbook], empirical work in this area is sparse. Much of this literature is based on the recognition that pure cost-plus regulation eliminates firms’ incentives to minimize costs or improve efficiency. While few regulatory processes are in fact purely cost-plus, there has been little analysis either of how close existing regulatory procedures are to such a system, or of the extent to which incentive-dampening effects of regulation induce higher costs. Gollop and Karlson (1978) examine the effects of automatic fuel adjustment mechanisms (FAM), and find little evidence of FAM-induced input biases. They do find some FAM-induced X-inefficiency, consistent with predictions of the theoretical models. Joskow and Schmalensee (1986) argue that two features of standard public utility ratemaking – investment prudence reviews and regulatory lag – distinguish regulation from a pure cost-plus contract. The design and adoption of explicit incentive policies, while increasing during recent years, has been somewhat arbitrary, and the effects of these policies are uncertain. As Joskow and Schmalensee point out, many of these schemes may introduce new distortions of firm behavior.

Regulation may also raise firms’ costs by increasing financial risk and the cost of capital. The nature of the regulatory process will affect the systematic risk faced by regulated firms and therefore their cost of capital [Brennan and Schwartz (1982)]. These effects may vary with economic conditions if regulatory constraints interact asymmetrically with variations in economic activity. Unfortunately, the likely dependence of the cost-of-capital on specific features of the regulatory process has not been generally recognized by regulatory agencies. This may be a serious problem as regulators consider changes in regulatory rules and procedures that reallocate risk between firms and consumers [Joskow and Schmalensee (1986)]. For example, Clarke (1980) analyzes the effects of fuel adjustment mechanisms on systematic risk, using firm level financial data and a comparative time-series approach. He finds that FAMs reduce regulated utilities’ systematic risk, but have little if any independent effect on risk-adjusted stock market values.

Norton (1985) analyzes risk effects of regulation for a sample of electric utilities operating in “strongly regulated”, “weakly regulated”, and “unregulated” states. He compares “betas” from a Capital Asset Pricing Model of returns, and concludes that regulation reduces systematic risk. Norton’s distinction between “strongly regulated” and “weakly regulated” states is based on differences in the resources devoted to regulation; unregulated firms are in states without commission regulation. This methodology raises two concerns. First, using regulatory
“inputs” to distinguish the intensity of regulation across states is arbitrary; characterizing the regulatory environment by its effects seems preferable [Hendricks (1975), Joskow (1974)]. Second, failing to control for differences in economic characteristics across utilities, particularly with respect to cyclical sensitivity of demand, may bias the results. In particular, utilities that depend on cyclical industries (such as steel, coal, and iron mining) for a large fraction of their revenues are likely to have different risk characteristics than will utilities serving largely residential, commercial, and less cyclical industrial loads. Norton’s six “unregulated” utilities, located in Minnesota and Texas, may be unusually risky, given their dependence on the highly cyclical iron ore and petrochemical industries.\footnote{In 1983, revenues from large industrial customers accounted for 50 percent of investor-owned utility (IOU) revenues in Minnesota, 36 percent in Texas, but only 28 percent for the United States as a whole [Edison Electric Institute (1983, p. 67)].}

Despite theoretical interest in franchise bidding as an alternative to traditional commission regulation [Demsetz (1968), Williamson (1976)], there has been relatively little empirical analysis of the consequences of using municipal franchising in place of commission regulation. Williamson’s (1976) case study of a cable TV (CATV) franchise illuminates some problems that arise in municipal franchising.\footnote{Research by Prager (1986) and Zupan (1987) suggests that the contractual problems Williamson identifies have not been very serious in practice, except perhaps in large urban areas franchised since 1980.} Shew (1984) compares the costs of CATV franchising requirements with subscriber benefits and concludes that the franchising process has led to a significant amount of wasteful expenditures. These arise as municipal authorities force potential franchisees to compete over the services they provide to the municipality, rather than on the basis of product price and quality. Rather than yielding optimal prices, municipal franchise bidding results in excessive expenditures on cable facilities that are valued highly by local politicians, but are of lower value to consumers.

6.2. Multi-firm regulation

Many studies examine the cost effects of economic regulation in multi-firm industries. Joskow (1981), Sloan and Steinwald (1980), Sloan (1981) and Melnick, Wheeler and Feldstein (1981) use an interstate comparative approach to evaluate the effects of rate regulation and/or certificate of need (entry) regulation on hospital costs. They generally conclude that rate regulation tends to reduce costs, but that certificate-of-need (entry) regulation does not.

Pre-deregulation studies of airline costs focused primarily on the consequences of quality competition (low load factors), and secondarily on the costs of
inefficient service to small communities and the effects of regulation on labor costs. Comparative time-series analyses including post-deregulation cost data suggest that other regulatory inefficiencies may have been much larger. There have been dramatic changes in airline route structures (hubbing), aircraft utilization, and labor productivity since deregulation [Bailey, Graham and Kaplan (1985, chs. 4, 5 and 8), Morrison and Winston (1986)]. These changes are convincingly attributed to airlines’ ability to optimize their routes free from CAB certification restrictions, as well as to competition. More work remains to be done in this area.

Cost effects of ICC regulation of surface freight transportation received considerable attention in the regulation era. Regulatory route restrictions led to considerable inefficiency in the trucking industry. Private carriers and exempt agricultural carriers inherently tend to have unbalanced loads, and regulations that prohibited them from hauling regulated commodities on their return trips made empty backhauls a severe problem [MacAvoy and Snow (1977, pp. 25–26)]. Route and commodity restrictions on regulated carriers (including many one-way authorities) increased their level of empty backhauls and partial loads relative to the unconstrained level; circuitous route authorities and gateway restrictions (which prohibited carriers from travelling via the most direct route) increased route mileage.

Despite widespread agreement on these qualitative effects of regulation, few studies attempt to quantify them. Moore (1975) estimates that unregulated carriers’ costs would decrease by $3.2 billion (in 1986 dollars) if their empty backhaul level were reduced to that of regulated carriers. Moore infers cost effects for regulated carriers indirectly, based on assumptions about regulatory price effects and calculations of rents to capital and labor. Combining Moore’s (1978) assumptions with a more plausible 10 percent discount rate to translate firms’ rents into annual terms implies cost inflation of 8–11 percent of revenues, or roughly $4.5 billion annually in 1986 dollars. These calculations could now be refined using data on deregulated system operations; we await such a study.

Investigations of regulatory cost increases in rail transportation have focused on two areas. The first area is ICC restrictions on route abandonments, which require railroads to maintain service on lightly travelled, unprofitable routes, and result in excessively large systems (from a cost-minimizing standpoint). Friedlaender and Spady (1981) analyze this issue in a structural model of railroad and trucking cost functions and transportation demand functions. Using partial equilibrium analysis, they find that a 10 percent reduction in low-density track would reduce costs by $1.1 billion in 1986 dollars (p. 134; see, however, the caveats on p. 142). Keeler (1983) summarizes various authors’ estimates of the

---

16Moore’s (1978) own calculations assume an after-tax discount rate of 35 percent (70 percent pre-tax).
total cost of excessive route mileage at $900 million to $1.8 billion in 1986 dollars.

A second major regulatory cost arises from inefficient freight car utilization. Studies typically conclude that the ICC set car rental rates (the rates that railroads pay for using other railroads' boxcars on their system and that shippers pay for keeping cars on their sidings) below their opportunity cost, resulting in too little investment in freight cars and suboptimal utilization rates. To counteract some of these effects, the American Association of Railroads established rules requiring cars to be returned via the most direct routing—which also contribute to inefficient utilization. Estimated annual costs of these inefficiencies range from $2.7 to $3.1 billion in 1986 dollars (Keeler (1983)).

Essentially all the empirical approaches discussed above also have been used to measure the effects of environmental, health, and safety regulations on production costs. Perl and Dunbar (1982), for example, use a simulation approach to estimate the effects of the New Source Performance Standards on the cost of producing electricity. Gollop and Roberts (1983) exploit variations across states in the intensity of environmental constraints on electric generating plants to estimate the costs of sulfur emissions constraints. Joskow and Rose (1985a, 1985b) make use of variations over time and space in power plant emissions scrubbing requirements, as well as variations in intensity of scrubbing, to estimate the costs of scrubbers. The general approaches to estimating the effects of regulation on costs can be applied quite widely to analyze the effects of "social" regulation as well as "economic" regulation.

7. The effects of regulation on innovation and productivity growth

Technological change and innovation has played a central role in increasing real incomes in the United States over time. Several heavily regulated industries have exhibited unusually high rates of productivity growth over long historical periods. These include the electric power industry until 1970 [Joskow and Rose (1985a, 1985b) and Joskow (1987)], the telecommunications industry, the airline industry, and the trucking industry. Others, for example railroads, have poorer productivity records [when correctly measured; see Caves, Christensen and Swanson (1981b)]. There are a number of channels through which price and entry regulation might affect incentives to innovate. Price regulation could change the pattern of expected returns to innovation. Shifting competition from price to non-price dimensions could increase incentives for rapid adoption of product innovations. Restrictions on entry and approval of rates for new services could delay the introduction and slow the diffusion of product, service and process innovations—both directly and indirectly, by raising the costs of introduction and diffusion and reducing the present value of net revenues associated with the
innovation [Braeutigam (1979)]. In spite of this, surprisingly little empirical research has been devoted to quantifying the effects of price and entry regulation on innovation and productivity growth.

The existing evidence on the effects of economic regulation on innovation includes anecdotes, case studies, and a few systematic econometric studies. Gellman (1971) documents several examples of how the Interstate Commerce Commission’s rate policies delayed the introduction of piggyback rail cars and the “Big John” rail car in the late 1950s and early 1960s. MacAvoy and Sloss (1967) argue persuasively that the adoption of the unit train was delayed considerably by ICC commodity rate restrictions. These conclusions are reinforced by the rapid increase in piggyback rail carriage, multiple car and unit trains after ICC deregulation in the mid- to late-1970s [MacDonald (1986)].

Phillips (1971) argues that CAB regulation of the airlines did not retard innovation, and there is casual evidence to suggest that CAB ratemaking policies encouraged rapid diffusion of larger, faster aircraft. Bailey, Graham and Kaplan (1985) document significant gains in airline productivity after deregulation, but these gains cannot be attributed specifically to an increased rate of technological innovation. Shepherd (1971) hypothesizes that regulation retarded innovation in telecommunications in a variety of different ways, but provides little empirical support for these hypotheses.

A few industry-specific econometric studies have tried to measure regulatory effects on productivity growth and innovation. Nelson and Wohar analyze productivity growth in the electric power industry using an A–J type of model and exploiting variations in the intensity of regulation. They find that regulation had both positive and negative effects on productivity growth, depending on the time period examined. Given the problems they have identifying a meaningful regulatory effect, however, these results should be interpreted cautiously. Joskow (1981) finds that rate and certificate-of-need regulation of hospitals slowed the diffusion of CT scans, and pushed them out of hospitals into physicians’ offices [see also Russell (1979)]. Caves et al. (1981a) find that productivity growth in the regulated U.S. railroad industry lagged substantially behind that in the unregulated Canadian railroad industry. During the 1956–1963 period, Canadian railroads averaged 1.7 percent productivity growth, versus 0.6 percent for U.S. railroads. Over 1963–1974, the differences were even more striking, at 4.0 percent versus 0.1 percent. They attribute these differences to the U.S. regulatory environment.

There has been considerably more interest in measuring the effects of environmental, health, and safety regulation on productivity growth and innovation. Peltzman (1973) uses a comparative time-series approach and a simulation approach to measure the effects on new drug introductions of the 1962 amendments tightening the FDA’s regulation of the safety and efficacy of prescription drugs. He finds that the costs of reduced innovation, as measured by the reduction in the number of new drugs introduced, greatly exceeds savings from
avoiding ineffective drugs; total costs, including the cost of reduced competition, are estimated at 5–10 percent of the annual $5 billion expended on drugs. Wiggins (1981, 1983) presents a related and updated analysis that disaggregates drugs into therapeutic categories [see also Grabowski and Vernon (1983) and Temin (1980)]. He finds that FDA regulations significantly reduced new drug introductions during the 1970s [Wiggins (1981)], and reduced company expenditures on research [Wiggins (1983)].

Christaiens and Haveman (1981) use crude measures of variations in the intensity of federal regulation over time to measure the aggregate effects of “public regulation” on productivity growth. They find that increased regulatory constraints are responsible for about 15 percent of the slowdown in productivity growth in manufacturing between 1973 and 1977 [see also Crandall (1981), Denison (1979), and Siegel (1979)]. The crude measures of regulation used and the almost perfect correlation between increases in regulatory intensity and other economic shocks over time (e.g., energy price increases, inflation, stagnant economic growth, import competition) limit the confidence one can place in these point estimates and suggest the desirability of further, more sophisticated analyses.

It is distressing that so little effort has been devoted to measuring the effects of regulation on innovation and productivity growth. Much of what we do know is now quite dated. The static gains and losses from regulation are probably small compared to the historical gains in welfare resulting from innovation and productivity growth. Further research on what, if any, effect regulation has on the dynamics of productivity growth and the development of new goods and services therefore seems essential.

8. The effects of regulation on product quality

Empirical analyses of regulatory effects on product quality have been fairly limited. The most intensively studied regulation–quality interaction has been in the airline industry, perhaps because we have a good theoretical model of the relationship between quality, price regulation, and the number of competing firms. Following Douglas and Miller (1975) [see also Schmalensee (1977)], the “quality” of airline service is measured by both the frequency of departures (the more departures, the more likely will there be a flight close to a passenger’s preferred departure time) and the probability of finding an available seat on the flight closest to the passenger’s preferred departure time. Empirical applications usually summarize both dimensions of quality by the average load factor, that is, total passengers divided by seats available on a route (load factors also may capture a third dimension of quality: expected crowding on a flight). Douglas and Miller (1975, chs. 2 and 6) also perform a more sophisticated stochastic simulation involving departures, flight size, and passenger valuations of time.
Most studies [see Douglas and Miller (1975), Keeler (1972), Eads (1975), Graham, Kaplan and Sibley (1983), Bailey, Graham and Kaplan (1985)] find that price regulation induces non-price service competition, yielding equilibria that on average give passengers too much quality; that is, given consumer valuations of service quality, flights are too frequent, load factors are on average too low, and costs are too high. Furthermore, the price/quality relationship depends on the number of competing firms on each route. Routes with large numbers of firms have very low load factors, as service quality competition drives average cost per passenger up to (or above) average revenue per passenger. In monopoly markets, nonprice competition does not occur, so passengers pay relatively high fares and get relatively low quality. In their empirical work, Douglas and Miller show that average load factors vary inversely with the number of airlines certificated to serve a particular route, as predicted by their theory of price-constrained competition. Despite the fact that entry was restricted, the industry did not appear to earn sustained excess profits, or even reasonably stable profits during the regulated era [Douglas and Miller (1975, p. 18), Bailey, Graham, and Kaplan (1985, pp. 23–26)]. Non-price competition, and perhaps supracompetitive labor costs resulting in part from regulation (discussed in Section 9), appear to have ensured that high airline prices did not lead to excess returns for the owners of airline firms.

Deregulation of the airline industry in 1978 provides data for a comparative time-series analysis of service quality. Graham, Kaplan and Sibley compare load factors after deregulation with those observed before deregulation. They find that load factors increased (as expected); based on the traditional models of airline regulation, this should reduce average service quality. However, increased use of peak-load pricing and withholding of high fare seats in anticipation of late reservations mitigated this apparent decline in service quality for time-sensitive passengers. Using a simulation approach, Bailey, Graham and Kaplan also show that convenience did not decrease appreciably after deregulation, despite the increase in average load factors. Morrison and Winston (1986) find that travel time fell in smaller markets and increased in larger markets (for an average increase of 5 percent), while flight frequencies increased in almost all sizes of markets (by 9 percent on average). They conclude that net service quality has improved with deregulation, as has aggregate consumer welfare. Increasing load factors, and hub-and-spoke networks, along with many other supply side changes after deregulation, also have helped to reduce the cost per passenger-mile by increasing productivity.

There is anecdotal evidence to suggest that trucking regulation had adverse effects on product variety, by foreclosing quality-varying rates. Shippers complained that they were disadvantaged by their inability to obtain low rate/low quality or high rate/high quality service [see MacAvoy and Snow (1977, pp. 10–14)]. We are aware of no studies that attempt to quantify these effects.
Joskow (1980) uses a cross-sectional interstate comparative approach to measure the effects of rate and certificate-of-need regulation on the service quality of hospitals. Hospital quality is inversely indexed by the probability that a patient will be turned away because the hospital is full. Joskow argues that the characteristics of hospital insurance and provider reimbursement systems in the 1970s gave hospitals incentives to engage in quality competition. He finds that hospitals located in states that regulated rates and entry had lower service quality than those in states that did not impose such regulations.

Munch and Smallwood (1980) examine the effects of a variety of regulations on the solvency of property/liability insurance firms, using an interstate comparative approach. Solvency is a quality attribute because policyholders prefer to be insured by a company that will be able to pay off if a loss is incurred. Munch and Smallwood find that the probability of insolvency is reduced by state regulations that impose minimal capital requirements on insurers. This reduction is accomplished by making entry at small scale more costly, thereby reducing the number of small entrants. They also find that firms operating in states with rate regulation have a lower probability of insolvency, but that the difference between states with prior approval regulation and those without it is not statistically significant. Frech and Samprone (1980) attempt to estimate the welfare consequences of non-price competition in regulated property insurance markets, but their method for quantifying the regulatory-induced increase in non-price competition leaves much to be desired.

Finally, we note that regulatory agencies have sometimes claimed that regulated electric, gas and telephone utilities systematically build systems with excessively high reliability (quality); in some recent cases, agencies have disallowed cost recovery on plant and equipment deemed to be in excess of “prudent” reserve requirements. Using the model developed by Joskow (1974), Carron and MacAvoy (1981) argue that regulatory quality effects are exactly the opposite. They expect high service quality during periods of increasing productivity and stable prices, when regulatory constraints do not bind. As inflation increases nominal costs and regulators resist price increases, firms reduce their capital investment and therefore their service quality. Carron and MacAvoy (1981, pp. 48–53) cite declining reserve capacities, increasing delays, and increasing equipment problems during the 1970s in support of their argument. While Carron and MacAvoy’s argument is intuitively appealing, the evidence they present is incomplete and not entirely convincing. The quality effects of natural monopoly regulation remain uncertain.

9. The distributional effects of regulation and deregulation

Distributional consequences of regulation play a fundamental role in explaining the incidence of regulation and regulatory change, according to modern political
economy theories. Alternative regulatory arrangements (including what is popularly called "deregulation") generally imply different distributions of benefits and costs. These distributional effects drive the competition of various interest groups, which in turn determines the nature of regulation through the political process.

Until recently, information on the distributional consequences of regulation was primarily a byproduct of the studies of regulatory price, cost, and quality effects that we have already discussed. In some sense, this is not terribly surprising. Most regulatory research draws on neoclassical economic theories, which focus on economic efficiency. Furthermore, tracing through the ultimate incidence of changes in prices and costs is often quite difficult. Identifying "first-order" winners and losers from regulation and measuring the magnitudes of these effects is, however, practicable; this is the focus of a growing segment of the regulation literature.

A number of authors have analyzed distributional effects of regulation in the context of implicit or explicit tests of political economy models of regulation. Kalt (1981) measures the winners and losers from 1970s petroleum price regulations by comparing the outcomes of regulation with the simulated outcomes in the absence of price regulation. These results are then used to analyze Congressional voting behavior. Kalt finds that both constituents' economic interests and measures of congressmen's ideology are important explanators of votes. A similar approach is used to study the 1977 Surface Mining Control and Reclamation Act (SMCRA) by Kalt (1983a) and Kalt and Zupan (1984). Kalt and Leone (1986) use a simulation approach to examine the effects on regional incomes of deregulation of natural gas prices. Olson and Trapani (1981) attempt to measure the effects of CAB regulation of airlines using a simulation approach that compares regulatory outcomes with various norms. They argue that consumers lost from regulation, aircraft manufacturers benefited, and airlines benefited during some time periods. These issues also have been analyzed for various "social regulations"; see Pashigian (1985) and Oster (1982), for example.

Rather than restrict attention to these tests of political economy models of regulation, we discuss a broad range of evidence on the distributional effects of regulation. We consider four types of regulatory redistributions: transfers to the owners of regulated firms (profits), transfers to factors of production such as labor ("rent-sharing"), transfers among consumer groups, and transfers among producers. Empirical studies of these effects tend to focus on regulated multi-firm industries; evidence from franchised monopoly regulation is discussed where available.

9.1. Profits

Positive theories of regulation predict that regulated firms should, at least in some cases, gain from regulation. This is particularly true where members of the
industry were strong advocates of regulatory intervention. As might be expected from the diversity of regulatory price effects, profit effects vary considerably across regulated industries.

A variety of approaches have been used to measure profitability effects. As discussed in Section 3, market values of regulatory assets can provide one of the cleanest tests of regulatory effects on profits. Kitch, Isaacson and Kasper (1971) provide one of the earliest applications of this approach in their study of Chicago taxicab regulation. They use taxicab medallion prices to estimate regulatory rents of $115 million (1986 dollars). Although this approach could be applied to estimate differences in taxicab regulation across municipalities, there has been little additional work in the area. This method has, however, been applied extensively to analyze trucking regulation. Moore's (1978) sample of 23 certificate sales suggests that certificate values are roughly 15 percent of gross revenues. Applying this to aggregate industry revenues in 1972 yields a present discounted value of rents of $5.5 to $7.9 billion in 1986 dollars. Breen's (1977) study of household goods carriers operating certificates and Frew's (1981) study of common carrier certificate values also find evidence of substantial certificate values. The sharp decline in certificate values around the period of trucking deregulation confirms the interpretation of certificate values as regulatory rents [Moore (1986), Mabley and Strack (1982)].

A second type of asset market approach, the event study technique, relies on changes in the regulatory environment to analyze regulatory effects. Rose (1985a) uses this approach to analyze the effects of ICC administrative reforms and congressional deregulation on the trucking industry. She finds substantial share price responses to deregulation, with market values of publicly traded general freight carriers declining by 15–19 percent and those of specialized commodity carriers declining by 9 percent. These correspond to a decline of $925 million (1986 dollars) in capitalized rents for the 32 firms in her study, or about 8.8 percent of 1978 gross revenues. Applying the sample rent/revenue ratio to aggregate revenues suggests that the present discounted value of rents earned by the 345 Class I general freight carriers in 1978 was $2.6 billion in 1986 dollars [Rose (1987)].

Levin (1981) uses a structural estimation/simulation approach to estimate regulatory effects on railroad profitability. Although the precise magnitude of effects depends on which competitive scenario is selected, he finds substantial increases in profitability under all but the marginal cost pricing cases. Net income under regulation was $570 million in 1986 dollars, or a 2 percent rate of return on replacement cost of assets. Under moderate competition and deregulation, estimated net income was $3.4 to $11.3 billion (1986 dollars). This suggests substantial transfers away from capital under railroad regulation.

Results for the airline industry are mixed. Few studies during the regulatory era specifically address the question of regulatory profits. A number of authors
argue that higher prices raised service quality, raising costs and preventing high prices from translating into high profits. While this may indicate that regulatory price policies failed to increase rates of return, regulatory entry policies may have nevertheless raised profits above normal levels [see Schmalensee (1977)]. Indirect support for this view is found in the strong opposition of most trunk carriers to airline deregulation [Kahn (1983)]. Opposite results are found by Morrison and Winston (1986), who find that simulated deregulated profits would have been higher than were 1977 regulated profits. They argue that observed declines in airline profitability in the early 1980s were largely the result of fuel price shocks and macroeconomic conditions, not deregulation.

Natural gas and petroleum price regulations, while not themselves creating rents, have generated enormous rent transfers to and from various interest groups. For example, Kalt (1981) estimates that crude oil price controls reduced the incomes of producers by $19 to $65 billion annually (in 1986 dollars) over 1975–1980, and increased the income of refiners by roughly 60 percent of this amount over the period. Smith, Bradley and Jarrell (1986) use an event study technique to estimate refiner gains from the early Crude Oil Allocation Program adopted in response to OPEC’s 1973 price increases; their results indicate substantial refiner gains.

9.2. Factor rent-sharing

Many formal models of the effects of regulation and most studies of regulatory effects on prices assume that factor prices are independent of the regulatory environment. This assumption may be invalid for many regulated industries. Regulation may transfer rents to other factors of production, even when capital earns normal returns. Empirical analyses of the effects of economic regulation on factor returns (other than capital) have focused almost exclusively on labor. There are several channels through which the regulatory process may alter the relative bargaining positions of regulated firms and workers. First, to the extent that the regulatory process allows wage increases to be quickly and completely passed through by higher prices, a firm’s incentives to be a tough bargainer are diminished, and higher wage settlements may result [Hendricks (1975), Ehrenberg (1979)]. Some forms of price regulation—such as the operating ratio constraint used by the Interstate Commerce Commission to evaluate trucking rates—may reduce the shadow cost of labor, perhaps exacerbating this tendency [see Daughety (1984), Moore (1978)]. We also note that gains by unionized employees may be realized, at least in part, by non-union workers as well. The labor economics literature suggests that union “threat effects”, among other factors, tend to raise non-union wages in industries with large union gains [see Lewis (1963, 1986)].
Second, entry restrictions can create a situation in which suppliers that are shut down by a strike cannot be replaced by rival suppliers, because the alternative suppliers are not authorized to provide services in the "struck" markets. This has conflicting effects on the distribution of bargaining power. On the one hand, it may increase union bargaining power by increasing the disruption caused by strikes and the consequent public pressures to settle them [Bailey, Graham and Kaplan (1985, pp. 96–97)]. On the other hand, eliminating the ability of potential competitors to take over a firm's customers during a strike is likely to reduce strike costs to the firm, other things equal, by eliminating post-strike customer defections. This tends to increase the firm's bargaining power vis-à-vis the union.

Third, to the extent that regulation restricts entry, it may be easier to organize a regulated industry and easier to sustain high wages without the threat of entry by lower cost non-union suppliers. Some authors have suggested that entry regulation may enable a union to cartelize an industry and realize monopoly profits for its members [Arnold (1970), Annable (1973)].

Finally, regulation may introduce political considerations into input choice decisions. This will apply to non-labor inputs as well as labor inputs. For example, utilities may be subject to political pressures to buy local products (e.g., coal), rather than cheaper substitutes from suppliers in other states, to bolster the local economy. Depending on the state, they might also come under pressure to use unionized employees even when they otherwise might not, or to sign lavish wage agreements [Ehrenberg (1979)].

A number of studies have investigated the effects of regulation on wages. Hendricks (1975) examines wage settlements by electric utilities in the context of three different models of the regulatory process, using an interstate comparative approach. He finds that regulated firms' bargaining incentives and wage settlements depend on the nature of the regulatory constraint they face. Firms operating under a non-binding rate-of-return restraint (due to regulatory lag or benign neglect) are more aggressive bargainers than are firms that expect the regulatory agency immediately to adjust rates to pass through higher wages. As a consequence, wages are lower for utilities that expect to bear increased costs themselves, and higher for utilities that expect regulators to flow through cost increases.

While Hendricks (1975) focuses on differential wage patterns within a regulated industry, most subsequent empirical work attempts to measure regulatory effects on average wage levels. The evidence on this question is mixed. Hendricks (1977) investigates the distribution of wages for workers across fourteen regulated industries and the unregulated manufacturing sector, using micro-data on individual workers to estimate a conventional human capital earnings equation with controls for occupation, industry concentration, and regulation. For most occupations and most regulated industries, regulation appears to have zero or negative
effects on average wage levels. The dominant exceptions are truck drivers in the trucking industry, electricians in radio and television, and possibly certain airline occupational categories; these groups are associated with higher wage levels than in manufacturing as a whole. Carrol and Ciscel's (1982) inter-industry study of executive compensation concludes that executive compensation is significantly lower in regulated industries (utilities and transportation) than in unregulated industries. This work suggests that regulatory wage gains may be limited to certain industries, and may be stronger for workers in certain key occupations (such as drivers in the trucking industry).

A number of other studies explore wages within a particular regulated industry. Ehrenberg's (1979) detailed empirical analysis of New York Telephone Company worker salaries suggests that New York Telephone paid higher wages to many categories of workers. Hendricks (1975, 1977), on the other hand, finds that average levels in the electric utility industry are below those for comparable workers in unregulated sectors. Substantial empirical effort has been focused on wage levels in regulated transportation industries. The relationship between pilots and regulated airlines has been studied extensively [Baitsell (1966), Kahn (1971), Pulsifer et al. (1975)]. The pilots' union has been successful in negotiating extremely attractive wage and work rule arrangements, which appear to have been at least partially a consequence of regulation [Bailey, Graham and Kaplan (1985, pp. 139–147)]. Hendricks, Feuille and Szerszen (1980) use micro data on individual workers and data on characteristics of collective bargaining agreements to investigate the extent of regulated airline workers' gains relative to a manufacturing benchmark. They find that airline workers across a wide variety of occupations have higher mean wages than their manufacturing counterparts, even after controlling for worker quality (see also Bailey, Graham and Kaplan, p. 18). This differential appears to be associated primarily (but not exclusively) with the high levels of unionization and concentration in the airline industry, which may themselves be functions of regulation. Industry responses to deregulation confirm positive regulatory wage effects. Cappelli (1985) reports that virtually all airlines obtained some form of union concessions after deregulation, with concessions concentrated among pilots. Mechanics, whose skills are easily transferable in and out of the airline industry, were least affected by concessions [see also Card (1986)].

Similar analyses have been performed for the trucking industry. Moore (1978) uses a variety of aggregate earnings data and micro data wage equations to estimate union rent-sharing in the regulated trucking industry. Using a 50 percent union wage premium, he calculates Teamster rents from regulation at $2.6 to $3.4 billion (in 1986 dollars). There are a variety of problems with inferences based on these simple calculations. In particular, unionized workers in most industries earn higher wages than comparable non-union workers; to attribute the entire union
wage premium to regulatory rents requires an explanation for this more pervasive phenomenon. There also are a variety of reasons why cross-sectional estimates probably overstate the level of union premia [Lewis (1986), Rose (1987)].

The deregulation of the trucking industry in the late 1970s and early 1980s provides variation that can be used to estimate the extent of labor rent-sharing more precisely. As in the airlines industry, extensive low-wage, non-union entry (combined with the 1981–82 recession) squeezed existing union carriers. This led to substantial unemployment for union drivers and ultimately resulted in considerable contract concessions by the Teamsters Union [Perry (1986), Rose (1987)]. Econometric analysis of industry wage behavior by Rose (1987) confirms the contract evidence [see also Hirsch (1986)]. Using micro data estimates of industry wage equations for the 1973–85 period, she finds that union wage premia decline from an average of 50 percent over non-union wages during the regulated period to less than 30 percent during the deregulation period. This decline reduces the trucking union premium to the level of the average blue-collar union premium for the economy as a whole. Implied aggregate union losses are $700 million to $1.3 billion per year, or roughly twice the estimated annual losses for owners of trucking firms. Contrary to models of non-union rent-sharing, Rose finds little evidence of non-union wage declines or rent spillovers to truck drivers outside the regulated trucking sector.

Although regulation did not create rents for owners of railroads, labor appears to have gained from regulation, particularly through enhanced union bargaining power. The ICC, for example, required very costly labor protection agreements as a prerequisite to merger approvals [Lieb (1984)]. Rail work rules are among the most restrictive in industry, and tend to enforce higher pay and higher labor requirements. While these cannot be attributed solely to regulation (strong rail unions might have evolved independently of regulation), regulation may contribute to their effect. In the wake of rail consolidations and route rationalizations following deregulation, labor appears to have made limited concessions in some firms, although these do not approach the concessions made by trucking and airline employees [Lieb 1984].

9.3. Transfers among customer groups

Economic regulation may have important distributional effects across customer groups. The multi-product nature of many regulated industries and political influences on pricing that act through regulatory procedures and appointment processes create a situation ripe with potential cross-subsidies; these frequently are exploited to benefit particular interest groups [Posner (1971)]. Studies of these effects tend to be qualitative. Quantification of inter-customer distributional effects, when available, tends to be an outgrowth of analyses focusing on efficient
price structures in regulated industries. These studies are discussed at length in Section 5; we summarize their implications for income transfers below.

Cross-subsidies – i.e. subsidization of some categories of customers by others – is a common theme throughout public utility regulation. For example, it is sometimes argued by industrial customers that the rate structure for electric utilities is skewed toward lower rates for residential and commercial users and higher rates for industrial users, relative to efficient rates. Similarly, implicit coordination of rate-setting between the Federal Communications Commission and state regulatory commissions resulted in telephone pricing procedures that elevated long-distance rates substantially above costs and competitive levels in order to subsidize below-cost prices for local service, and raised urban rates to subsidize rural customers [Johnson (1982), Noll (1985a), MacAvoy and Robinson (1983), Bailey (1986)].

The ability of regulators to maintain such redistributive policies depends critically upon their control over substitute products or suppliers. In the case of telephone pricing, technological advances in microwave communications, combined with court rulings that eliminated AT&T's legal monopoly on long-distance service, created the possibility that high-price long-distance customers would leave the system, forcing price increases on the remaining customers. The subsequent reductions in long-distance prices have decreased, though not eliminated, subsidies of local service; regulators continue to resist rate increases in local service sufficient to cover its appropriable costs [Noll (1985a)].

As noted earlier, the airline fare structure under CAB regulation built in a variety of cross-subsidies. Fares on dense, long-distance routes were elevated substantially above costs, in part to generate profits that could be used to balance below-cost fares in sparse, short-haul markets. Congress and the CAB both were explicit about the protection of air service to small cities. Not only was service deemed essential, but service at “low” fares (relative to costs) was imposed. This was accomplished through both direct subsidies to carriers serving these markets and cross-subsidies of carriers with extensive route networks (enforced by CAB disapproval of carrier abandonment of unprofitable small city routes). Deregulation of prices, entry, and exit has increased fares in low-density, short-haul markets, eliminated air service to a number of small communities, and generally increased service frequency for small cities that have retained air service [Morrison and Winston (1986)]. Rail regulation created similar protection for shippers on low-density routes. Despite costs that appear to have been in excess of service value on many of these routes, the ICC refused to allow railroads to abandon the service, preferring instead to subsidize such service through transfers from capital owners.

Energy regulation created transfers among numerous special interest groups. Kalt (1981) analyzes these effects for petroleum regulation during the 1970s. He finds that regulation established many groups whose energy purchases were
heavily subsidized. Regulation-induced shortages gave rise to "priority" consumers, who were given "rights" to more certain production flows, and in some cases, lower prices. Agricultural users, for example, were given high priority, assured supplies; automobile users, low priority, uncertain, and often inadequate supplies [Kalt (1983b)]. Natural gas regulation created similar disparities among customer groups. Rather than increase natural gas prices to avoid chronic shortages during the mid-1970s, regulators relied on rationing. This led to a series of curtailment rules, under which industrial customers typically were shut off the system during shortages, existing residential customers were entitled to unlimited supplies, and new residential customers were excluded from hooking up to the system [Breyer (1982)].

9.4. Transfers among producer groups

Industries are not monolithic, and regulation may benefit some segments of an industry more than (or at the expense of) other segments. The intra-industry distributional effects of regulation have been studied most extensively for social regulation. Pashigian (1984) finds that environmental regulations tend to benefit large firms relative to small firms within an industry. Oster (1982) finds that state generic drug substitution regulations can be explained at least in part by the regulations' differential effects on two groups of firms: large pharmaceutical companies specializing in R&D and patented drugs, and smaller manufacturers specializing in generic drug production. Bartel and Thomas (1986) find that OSHA and EPA regulations have important intra-industry competitive effects that firms may exploit via "predatory" advocacy of particular regulations. There has, however, been little work that has attempted to quantify intra-industry transfers resulting from economic regulation.

In the case of the trucking industry, ICC regulation protected regulated carriers at the expense of exempt or partially exempt carriers. Rules restricting contract carriers to no more than eight shippers limited the size and expansion possibilities of contract carriers. Prohibiting private carriers from using owner-operators or sub-leasing their equipment and drivers on return trips to avoid empty backhauls increased their costs vis-à-vis regulated carriers. Owner-operators were limited to exempt commodities, or required to sign long-term contracts with regulated carriers. Contract and private carriers' dissatisfaction with ICC regulation in the late 1970s led to a conflict within the American Trucking Associations (ATA) over the ATA's lobbying position on deregulation initiatives, and ultimately may have contributed to the success of deregulation initiatives.

Similar intra-industry disagreements may have been operative in airline regulation. Kahn (1983) argues that a number of air carriers believed that regulatory route restrictions constrained their business opportunities by more than it con-
strained their competitors. These included Pan Am, which was confined to overseas operations by the CAB, and a number of intra-state and commuter carriers, which were precluded from expanding service by the CAB’s restrictive entry policy and long-standing refusal to certificate new trunk service. In contrast to the major trunks, these airlines basically supported deregulation.

Energy price and allocation regulations appear to have had substantial distributional effects across producers. Inter-state regulation of natural gas prices resulted in large transfers away from inter-state producers relative to exempt intra-state producers. This situation changed with enactment of the Natural Gas Policy Act of 1978 (NGPA), which extended the Federal Energy Regulatory Commission’s jurisdiction to intra-state wellhead prices. The NGPA’s complicated set of prices based on different categories of natural gas resulted in benefits to some wellhead producers and losses to others [Braeutigam and Hubbard (1986)].

As discussed earlier, Kalt (1981) finds that crude oil price regulations created an enormous pool of rents, which regulators divided among various producer and consumer interest groups. Initial regulations in 1973 resulted in large rents for refiners with historical ties to “old”, price-controlled oil. This system was replaced by the Entitlements Program, which redistributed income among refiners. This tended to raise the income of refiners who depended largely on imported crude oil, and lower the income of refiners who had access to domestic crude. In addition, a number of arbitrary redistributions, such as the “Small Refiner Bias”, were built into the system [Kalt (1981, 1983b)]. Smith, Bradley and Jarrell (1986) combine an event study with an economic model of oil regulation to analyze the joint effects of the OPEC price hike and U.S. oil regulations on producers. Their results, like Kalt’s, suggest that prior to the entitlements program, crude oil price regulation created large transfers from U.S. crude producers to those U.S. refiners with substantial access to price-controlled crude oil.

10. Conclusions

Systematic empirical analysis of the effects of economic regulation originated with Stigler and Friedland’s 1962 paper, which sought to measure the effects of state commission regulation of franchised electric utilities. Stigler and Friedland found that commission regulation had little or no effect on electricity prices. Since 1962 there have been several hundred scholarly studies of the effects of economic regulation. These have analyzed a broad range of industries, measuring regulatory effects against a number of different benchmarks, and using different types of data and a variety of empirical methodologies. This empirical analysis
has coincided with the development of both positive and normative theories of regulation and its effects.

The empirical regulation literature of the last twenty-five years clearly demonstrates that regulation frequently has substantial impacts on the behavior and performance of regulated firms. It is, however, impossible to generalize simple propositions about the effects of economic regulation; we cannot, for example, conclude that economic regulation always leads to lower prices than would emerge in the absence of regulation. The nature and magnitude of regulatory effects vary substantially, depending on the structure of the regulatory process, the industry being examined, and the economic environment. The diversity of observed regulatory effects should not be surprising. The term “economic regulation” covers many different types of economic control applied to quite diverse industries with a variety of objectives. Several common themes emerge from the empirical analyses, however.

1. The effects of economic regulation often differ considerably from the predictions of “public interest” models, which presume that regulation is intended to ameliorate market imperfections and enhance efficiency. This conclusion follows not simply from the observation that regulation is the outcome of a political process, but from analyses of the impacts of regulatory intervention.

2. In classical “public utility” industries, price regulation generally constrains prices below the level an unconstrained monopolist with a legal exclusive franchise would choose. The structure of prices and distribution of revenues across classes of customers often reflect distributional and political objectives, however, rather than efficiency objectives. Furthermore, regulated prices may not be lower than prices would be under a fundamentally different industry structure with multiple firms and free entry.

3. In regulated markets with multiple competing firms, the effects of entry regulation on prices are more complex. In some industries (airlines, surface freight transportation, insurance), price and entry regulation seem to have been introduced to protect incumbents for competition. Despite this intent, the ability of price regulation to transfer income to the owners of the regulated firms has been sharply constrained by non-price competition and factor rent-sharing. Institutional inertia further limits the benefits of regulation through time, as regulated firms are constrained from adapting to the introduction of competing products and suppliers. The immutability of regulation can transform protectionism into strangulation, as in the railroad industry. Regulation often persists despite the apparent absence of economic rents because regulation has so distorted industry structures that changes would lead to large losses for incumbents [McCormick, Shugart and Tollison (1984)].

In other multi-firm industries, price regulation has been introduced primarily to protect consumers from precipitous price increases. This is true of natural gas, petroleum, hospitals, and to some extent electric utilities with expensive nuclear plants. These regulatory initiatives appear to be self-limiting. At some point,
efforts to keep prices below market-clearing levels cannot be sustained. With prices below the marginal cost of additional supplies, shortages develop and the quality of service deteriorates. These effects generate intense pressures for regulatory reform.

(4) Economic regulation has important direct and indirect effects on the costs of production and the quality of service. Regulatory influences on input choices, $X$-inefficiency, and technological change tend to increase costs. Regulation also alters the quality and variety of services, although these effects often are difficult to quantify. It tends to increase service quality through non-price competition when regulated prices in structurally competitive industries are above competitive levels. Regulation may lower service quality when its intention is to keep prices below their market-clearing levels.

(5) Simplistic “producer capture” models, which view regulation as a cartelization device by which firms transfer income from consumers to producers, are just that—too simplistic. The distributional impacts of regulation are complex and vary with economic conditions and across industries. Moreover, labor appears to be an important beneficiary in a number of industries; perhaps more so than owners of regulated firms. Price and entry regulation seem to be especially conducive to the development of strong unions that can use their bargaining power to exact higher wages.

(6) Although the performance of regulated firms is sensitive to prevailing economic conditions, regulatory structures are quite impervious to exogenous economic forces. Regulatory systems tend to respond only to profound changes in the economic and political environment. The massive economic disruptions of the 1930s gave rise to a vast array of federal regulations, most of which persisted through the next forty years. The recent wave of federal regulatory reforms arose from the substantial supply shocks and macroeconomic disturbances of the 1970s, which have been characterized as the most severe disruptions since the 1930s. These reforms have dismantled or refigured much of the 1930s federal regulatory apparatus.

Our understanding of the effects of economic regulation has advanced considerably over the last twenty-five years, but many questions remain unresolved. The profound changes in both regulatory institutions and economic conditions during the past decade provide a valuable opportunity to answer some of these questions, through careful analysis of the effects of these changes. A number of recent studies have measured early responses to regulatory reforms. In some cases, however, these may be observing transition behavior, rather than a steady-state response to a new regulatory environment. Further analysis of the behavior and performance of industries that have experienced major changes in regulation will be invaluable in discerning permanent impacts.

With the decline of U.S. federal regulatory efforts, the research payoffs to more intensive study of state regulation and to comparative studies of industrial performance in a variety of regulatory and ownership settings may be substantial.
State regulation of some industries has expanded to fill the void created by federal deregulation. These regulations are seldom uniform across states; growing experimentation and diversity at the state level provides valuable variation through which we can measure the effects of regulation. This research will proceed most productively if we improve our measurement of differences in the intensity and types of regulation across states, as well as developing better controls for differences in the economic characteristics of regulated firms. Inter-county comparative studies are more difficult, but as other developed economies consider “privatization”, deregulation, and regulatory reforms, there is probably much to be learned by comparing the outcomes of different approaches to industrial ownership and control.

Finally, the large collection of empirical measurements of regulatory effects developed over the last twenty-five years provides a data base to better distinguish among competing theories of the political economy of regulation. Much of the research that tests alternative positive theories of regulation has focused on legislative voting behavior (almost exclusively at the federal level), and particularly on the relationship between legislative voting behavior and constituent interests. This work is interesting and important. We do not find it surprising, however, that Congressional voting behavior reflects constituent interests. Nor do we find it surprising that the discretion of regulatory agencies is sharply constrained by political considerations. This tells us simply that regulation has effects on various economic variables, that these effects have distributional impacts that create constituent interests, and that groups representing diverse interests respond in the political arena. Interest group politics is not, however, per se inconsistent with a “public interest” view of regulation (whatever that means) or with competing general “private interest” theories. The work on the political economy of regulation must inevitably be carefully related to the effects of economic regulation and the way economic regulation is accomplished. The politics and economic consequences of regulation are intertwined in complex ways. Further effort to fold more traditional analysis of the effects of economic regulation into analyses of the political economy of regulation seems essential.

References


Ch. 25: Effects of Economic Regulation


McKay, D. (1976) 'Has the A-J effect been empirically verified?'. Social science working paper 132, California Institute of Technology.


Rose, N.L. (1985b) 'Unionization and regulation: The division of rents in the trucking industry', working paper no. 1684-85, Sloan School of Management, MIT.


Ch. 25: Effects of Economic Regulation


Siegel, R. (1979) 'Why has productivity slowed down?', *Data Resources Review*, 1:1.59–1.65.


