ECONOMIC REGULATION

Edited by Paul L. Joskow

CRITICAL IDEAS IN ECONOMICS 1

Economic Regulation

Edited by

Paul L. Joskow

Elizabeth and James Killian Professor of Economics and Management
Massachusetts Institute of Technology, USA

CRITICAL IDEAS IN ECONOMICS

An Elgar Reference Collection
Cheltenham, UK • Northampton, MA, USA

Published by
Edward Elgar Publishing Limited
Glensanda House
Montpellier Parade
Cheltenham
Glos GL50 1UA
UK

Edward Elgar Publishing, Inc.
136 West Street, Suite 202
Northampton
Massachusetts 01060
USA

A catalogue record for this book is available from the British Library.

Library of Congress Cataloguing in Publication Data

Economic regulation / edited by Paul L. Joskow.
— (Critical ideas in economics) (An Elgar reference collection)
Includes bibliographical references and index.
1. Trade regulation. 2. Public utilities—Law and legislation. I. Joskow, Paul L.
II. Series. III. Series: An Elgar reference collection.

HD3612.E265 2000
338.9—dc21 99–052947

ISBN 1 85898 947 7
Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

Acknowledgements

The editor and publishers wish to thank the authors and the following publishers who have kindly given permission for the use of copyright material.

American Economic Association for articles: William J. Baumol and David F. Bradford (1970), 'Optimal Departures From Marginal Cost Pricing', *American Economic Review*, **LX**, 265–83; William J. Baumol (1977), 'On the Proper Cost Tests for Natural Monopoly in a Multiproduct Industry', *American Economic Review*, **67** (5), December, 809–22; Dennis W. Carlton (1977), 'Peak Load Pricing with Stochastic Demand', *American Economic Review*, **67** (5), December, 1006–10; Tracy R. Lewis and David E.M. Sappington (1988), 'Regulating a Monopolist with Unknown Demand', *American Economic Review*, **78** (5), December, 986–98.

Blackwell Publishers Ltd for article: Mark Armstrong, Chris Doyle and John Vickers (1996), 'The Access Pricing Problem: A Synthesis', *Journal of Industrial Economics*, **XLIV** (2), June, 131–50.

Brookings Institution Press for articles: Sam Peltzman (1989), 'The Economic Theory of Regulation after a Decade of Deregulation' and 'Comments and Discussion' by Michael E. Levine and Roger G. Noll, *Brookings Papers on Economic Activity*, 1–41, 42–8, 48–59; Paul L. Joskow (1989), 'Regulatory Failure, Regulatory Reform, and Structural Change in the Electrical Power Industry', *Brookings Papers on Economic Activity*, 125–99.

Econometric Society for article: David P. Baron and Roger B. Myerson (1982), 'Regulating a Monopolist with Unknown Costs', *Econometrica*, **50** (4), July, 911–30.

Elsevier Science Ltd for article: Jean-Jacques Laffont and Jean Tirole (1990), 'The Regulation of Multiproduct Firms Part I: Theory', *Journal of Public Economics*, **43**, 1–36.

Journal of Law and Economics, University of Chicago Law School for article: Paul L. Joskow (1974), 'Inflation and Environmental Concern: Structural Change in the Process of Public Utility Price Regulation', *Journal of Law and Economics*, **XVII** (2), October, 291–327.

Kluwer Academic Publishers for articles: Donald J. Kridel, David E.M. Sappington and Dennis L. Weisman (1996), 'The Effects of Incentive Regulation in the Telecommunications Industry: A Survey', *Journal of Regulatory Economics*, **9** (3), May, 269–306; Jean-Jacques Laffont and Jean Tirole (1996), 'Creating Competition Through Interconnection: Theory and Practice', *Journal of Regulatory Economics*, **10** (3), November, 227–56.

Oxford University Press for articles: Mathew D. McCubbins, Roger G. Noll and Barry R. Weingast (1987), 'Administrative Procedures as Instruments of Political Control', *Journal of Law, Economics, and Organization*, **3** (2), Fall, 243–77; John Vickers (1997), 'Regulation, Competition, and the Structure of Prices', *Oxford Review of Economic Policy*, **13** (1), Spring, 15–26.

RAND for articles: William J. Baumol and Alvin K. Klevorick (1970), 'Input Choices and Rate-of-Return Regulation: An Overview of the Discussion', *Bell Journal of Economics and Management Science*, **1** (2), Autumn, 162–90; George J. Stigler (1971), 'The Theory of Economic Regulation', *Bell Journal of Economics and Management Science*, **2** (1), Spring, 3–21; Richard A. Posner (1971), 'Taxation by Regulation', *Bell Journal of Economics and Management Science*, **2** (1), Spring, 22–50; Leonard J. Mirman and David Sibley (1980), 'Optimal Nonlinear Prices for Multiproduct Monopolies', *Bell Journal of Economics*, **11** (2), Autumn, 659–70; Nancy L. Rose (1985), 'The Incidence of Regulatory Rents in the Motor Carrier Industry', *Rand Journal of Economics*, **16** (3), Autumn, 299–318; Richard Schmalensee (1989), 'Good Regulatory Regimes', *Rand Journal of Economics*, **20** (3), Autumn, 417–36.

University of Chicago Press for articles: M. Boiteux (1960), 'Peak-Load Pricing', *Journal of Business*, **XXXIII** (2), April, 157–79; Jean-Jacques Laffont and Jean Tirole (1986), 'Using Cost Observation to Regulate Firms', *Journal of Political Economy*, **94** (3), Part 1, June, 614–41; Nancy L. Rose (1987), 'Labor Rent Sharing and Regulation: Evidence from the Trucking Industry', *Journal of Political Economy*, **95** (6), December, 1146–78; Sam Peltzman (1993), 'George Stigler's Contribution to the Economic Analysis of Regulation', *Journal of Political Economy*, **101** (5), October, 818–32.

Yale Law School for article: Paul L. Joskow and Richard Schmalensee (1986), 'Incentive Regulation For Electric Utilities', *Yale Journal on Regulation*, **4** (1), Fall, 1–49.

Every effort has been made to trace all the copyright holders but if any have been inadvertently overlooked the publishers will be pleased to make the necessary arrangement at the first opportunity.

In addition the publishers wish to thank the Library of the London School of Economics and Political Science and the Marshall Library, Cambridge for their assistance in obtaining these articles.

Introduction

Paul L. Joskow

Economical and reliable supplies of a variety of basic 'infrastructure' services such as electricity, water, telecommunications and various transportation services play a critical role in supporting economic growth and development. Many of these infrastructure services have generally been viewed as natural monopolies and supply responsibilities have often been given to state-owned or privately owned monopolies whose prices, investment programs, labor policies, rules governing entry of competing suppliers and service quality are subject to government control. Consumers purchase services from these industries pursuant to regulated price schedules or 'tariffs' approved by government regulatory authorities rather than being determined by market forces. We refer to this type of government regulation of firms and industries as 'economic regulation' to distinguish it from other forms of government regulation of business firms (for example environmental, health and workplace safety regulation).

In an effort to respond to the poor performance record of these monopoly infrastructure sectors, the changing economic stakes and political power of interest groups involved in these sectors, as well as opportunities for competition to replace monopoly created by technological change, many countries have implemented or are contemplating major industry restructuring and regulatory reform programs for these sectors. These reform programs often involve restructuring these sectors by shifting supply responsibility to private enterprises (privatization) and relying more on competition and associated market mechanisms, rather than price and entry regulation, to allocate resources to and within potentially competitive segments of these infrastructure sectors (competition).

However, in most of the regulated monopoly infrastructure sectors subject to these types of reforms, certain important segments (for example electric transmission and distribution networks, gas and water distribution networks and perhaps local telephone networks) continue to be natural and/or legal monopolies (or duopolies) that require continuing regulation. Moreover, open and non-discriminatory access by new competitors to the network facilities controlled by these monopolies is necessary for effective competition to flourish in the newly competitive segments of these sectors. Accordingly, government regulators must play an important role in defining and enforcing the terms and conditions of competitive access to these networks. In addition, governments often retain a continuing role in designing the institutions that govern the newly created competitive segments, in monitoring their performance, and stand ready to impose new regulations if performance is unsatisfactory. These considerations imply that describing the recent wave of reforms in these sectors simply as 'deregulation' can be quite misleading. As these sectors are privatized and restructured, important segments remain monopolies that must be subject to regulation. Indeed, introducing competition into some segments that require access to the remaining monopoly segments can significantly complicate the regulatory problems that must be addressed and increase the costs of regulatory errors. Thus, the success of monopoly infrastructure sector reform depends in part on creating effective

regulatory institutions to govern the remaining monopoly segments of these sectors. As a result, the privatization and deregulation movement has, perhaps ironically, simultaneously led to an increased interest in understanding the strengths and weaknesses of alternative regulatory mechanisms and to an enormous amount of research devoted to the creation of 'incentive regulation' or 'performance-based regulation' as an alternative to the traditional 'cost plus' or 'cost of service' regulatory mechanisms that regulators have arguably relied upon primarily in the past.

This collection of readings includes both theoretical and empirical articles that discuss both normative aspects of economic regulation – how to regulate in a way that maximizes the net social value of the goods and services produced by firms subject to economic regulation – and articles that discuss the positive aspects of regulation – the forces that lead to the introduction, evolution and reform of economic regulation into certain sectors and the performance of economic regulation once it has been introduced. There is an enormous body of theoretical and empirical literature on these subjects. In this collection I have endeavored to include articles that represent fundamental contributions to this literature and span the range of topics that it covers. Recent texts by Armstrong, Cowan and Vickers (1994) and Laffont and Tirole (1993) cover many of these subjects in greater detail. A related set of readings appears in Bailey and Rothenberg-Pack (1995) and focuses on privatization.

Part I: Theories of Economic Regulation

The standard textbook rationale for subjecting an industry to price and entry regulation is that the industry is a 'natural monopoly'. A natural monopoly is an industry where supply costs have the characteristics that it is less costly to supply output in a single firm than in multiple competing firms. In a single product industry, a sufficient condition for natural monopoly is the presence of increasing returns to scale or, equivalently, economies of scale over the range of output defined by the aggregate demand for the product produced by the industry. As Baumol (Part II, Chapter 6, p. 185) discusses, in the case of firms that produce multiple products, the conditions for the supply technology to lead to the conclusion that supply has natural monopoly characteristics is more complex. In general, there need to be economies associated with joint production of the multiple products (economies of scope) and economies of scale associated with increasing the scale of production of one or more of these products. In addition, in the absence of government restrictions on entry, for a single incumbent firm with economies of scale to be able to exploit its position by charging prices higher than its costs of production, the supply technology must be characterized by the presence of sunk costs which allow an incumbent monopolist to deter entry of competitors when it raises prices above the level that covers its total production costs. That is, the markets served by the natural monopoly must not be easily 'contestable' by new entrants if prices rise above the incumbent's costs of production (articles on the theory of contestable markets can be found in Bailey and Rothenberg-Pack, 1995) and there must not be close substitutes for the products supplied by the natural monopolies (for example competition faced by railroads to transport freight from trucks and barges). In fact, industries like telephone networks, electric power networks, gas pipeline networks, railroads, and so on, which have traditionally been subject to public regulation, are very capital intensive and, once sunk, assets cannot be easily redeployed either to other locations or to other

production sectors. However, technological changes in these sectors over time have reduced the significance of economies of scale, economies of scope, and/or economies of vertical integration and increased opportunities to replace regulated monopoly in some segments of these industries with competitive markets. Moreover, because the institution of regulated monopoly dulls incentives for efficiency and increases the costs of political manipulation of the sector, the promise of cost savings resulting from economies of scale and economies of scope may be easily overwhelmed by the realities of the inefficiencies of regulated monopoly.

Economists writing in the late 19th and early 20th centuries argued that industries whose technology was characterized by economies of scale and whose cost structure was characterized by a high proportion of fixed to variable costs were unsuitable for sustaining competition. They argued further that industries with these cost attributes would 'naturally' evolve into monopolies and that the competitive process that led to monopoly outcomes would be characterized by wars of attrition with inefficient entry and duplication of facilities, price instability and service quality problems. They argued that the 'public interest' would be served by granting monopoly franchises to suppliers in these sectors and subjecting their price and non-price behavior to public regulation. In the US, these franchised monopolies were initially subject to regulation by local municipalities through franchise contracts, but regulation gradually moved to 'independent' state and federal regulatory agencies with specific responsibilities to regulate prices levels, price structures and other aspects of firm behavior in designated industries. These agencies in turn adopted 'cost of service' regulatory mechanisms that, in theory, allowed firms to set prices at a level sufficient to cover their 'reasonable' costs of providing service, including a 'fair rate of return' on the investments that they had made to fulfill their service obligation. In addition, laws and regulatory rules imposed an 'obligation to serve' on franchised monopolies that, among other things, required them to have adequate capacity to serve the needs of all consumers in their franchise areas economically and reliably based on price and non-price terms and conditions specified in their regulated tariffs.

In reality, few of the industries that evolved into public or private monopolies subject to economic regulation attained their monopoly position 'naturally' as a consequence of the dynamics of market forces. Instead, these firms generally have formal or de facto *legal* monopolies as a consequence of laws and regulations created and enforced by governments which restrict entry of competing suppliers into their markets. Moreover, these firms have sometimes been able effectively to extend their monopolies, and the expanse of regulation, by integrating into potential competitive segments (for example the supply of telephone and network equipment, the production of natural gas and the generation of electricity) to take advantage of real or imagined economies of vertical integration and economies of scope. In other cases, price and entry regulation has been imposed on industries where it is hard to imagine that there could be a natural monopoly problem (for example trucking and bus service). Moreover, regulatory practice is often inconsistent with the pursuit of widely accepted 'public interest' social welfare goals and the effects of regulatory practices on prices, costs and wages, the innovations suggest that regulation can lead to significant inefficiencies that impose higher costs on consumers.

Tariff structures, service connection and supply expansion policies have also been used extensively to engage in redistributive policies (for example subsidization and cross-subsidization), benefiting some interest groups and burdening others. For example, in telecommunications, it is almost universally the case that long distance and international calling prices have been kept

high (far above marginal cost) so that these services would generate net revenues in order to make a substantial contribution to the fixed costs of the local network, allowing the prices of local service to be kept low. This policy has been justified as promoting 'universal service', something that can be viewed as a desirable policy based on assumed network externality problems (the value of the network increases the more people a user can reach over it) or simply on the broader social desirability of giving all of a country's citizens access to basic infrastructure services at 'affordable' prices, but in reality is often the result of well-organized interest groups that are pursuing less admirable goals. Most countries have provided direct and indirect subsidies to electricity and telephone services supplied to farms, residences and small businesses in rural areas. Redistributive goals are also reflected as well in regional or national policies to favor local fuel suppliers, primarily coal (for example in the electric power sectors in England and Wales, and Germany) and to favor domestic equipment manufacturers and construction companies.

The actual behavior of regulatory agencies and the effects they have on firm and industry performance are incompatible with a positive theory of regulation that is based on the assumption that regulatory agencies pursue the 'public interest' as measured by standard yardsticks of social welfare or the 'natural monopoly' rationale for price and entry regulation (Posner, 1974). The conflict between the normative theory of what regulators 'should do' if they pursued standard social welfare goals and what regulators actually do in practice and the resulting effects of their behavior, has led to the development of a rich literature that examines the political economy of regulation and regulatory reform. This literature, which has evolved in conjunction with related work on positive political economy and public choice by economists and political scientists, examines the introduction, practices, consequences and evolution of regulatory institutions from the perspective of interest group politics and the structure of regulatory, legislative, executive branch and judicial institutions (Noll, 1989).

Part I contains a set of classic articles that have had an important impact on the development of positive theories of economic regulation. They focus on regulation as a political instrument for redistributing income rather than as an instrument for ameliorating market imperfections and increasing the efficiency with which goods and services are supplied and priced when an industry has natural monopoly characteristics. George Stigler's (1971) paper (Chapter 1, p. 3) is a seminal work that is one of the foundations for the last 25 years of research in economics and political science that focuses on the role of interests groups in the development of economic and regulatory policies. From Stigler's perspective, interests groups demand regulatory actions by government, not to pursue 'public interest' goals, but rather to get the government to impose regulations that benefit these groups at the expense of others. Politicians in turn supply regulatory actions, not to ameliorate market imperfections associated with natural monopoly or to increase aggregate social welfare, but rather to respond to the desires of interests groups that are in a position to deliver benefits to the politicians, for example increasing their probability of being reelected. The spread and character of regulatory actions reflects the interests of those groups that are in the best position to influence politicians.

Richard Posner's (1971) paper (Chapter 2, p. 22) focuses on the tendency for regulators to use the institution of regulated monopoly as an instrument of public finance to 'tax' some groups buying service from the regulated monopoly in order to subsidize other groups of consumers or input suppliers; what Posner refers to as taxation by regulation. The institution of regulated monopoly is an effective (though not necessarily efficient) vehicle for engaging in

such policies for two reasons. First, monopoly status makes it impossible (or at least extremely difficult) for competitors to respond to and undermine the viability of the price increases required to finance direct and indirect subsidy programs. Second, the magnitude and nature of the subsidies and cross-subsidies are buried in complex tariff-setting policies and resource-acquisition programs that are heavily insulated from meaningful public scrutiny. This system of hidden taxes and subsidies makes it easier for the interest groups that benefit from them to sustain them because they are more readily hidden from broader public scrutiny than are more transparent 'on budget' legislative tax and direct subsidy programs.

Posner's paper is particularly relevant today as these subsidies and taxes are revealed in the context of industry restructuring and the introduction of competition. Posner points out that competition undermines the ability to use regulated monopoly sectors as vehicles for taxing and subsidizing. Indeed, in some infrastructure sectors, in particular in telecommunications and freight transportation, the cross-subsidies reflected in regulated tariffs have themselves been a major stimulus for competitive entry as potential entrants saw business opportunities as long as they could compete to supply services in segments where regulated prices were higher than the associated stand-alone costs of supply in order to produce revenue that could be used to provide subsidies in other segments. The kinds of direct and indirect subsidies that exist today will not be sustainable in the competitive segments of the infrastructure segments (that is electricity generation, commodity natural gas and long distance telephone service) and, to the extent that governments try to sustain them as they now exist in the face of competitive entry, the effect will be to distort the nature and direction of competition.

The (1997) paper by McCubbins, Noll and Weingast (Chapter 4, p. 88) applies modern work in positive political economy that is based on rational actor models of legislative institutions and legislative behavior to model structurally the linkages between legislative politics and regulatory behavior. In this work, regulatory agencies are never truly 'independent', but rather are closely tethered to the wishes of key legislators with a strong interest in the effects of the regulator's actions. As the interests of those who control the relevant legislative committees changes so too does the behavior of the regulatory agencies that the legislature oversees. This paper provides the foundation for a growing body of research that expands models that focus on legislative control to include the influence of the executive branch and the courts in the analysis of the political and economic forces that influence regulatory agency behavior.

Joskow's (1974) paper (Chapter 3, p. 51) examines how the intensity of regulation responds to changes in the economic, technological and political environment faced by electric utilities in the US in the 25 years following World War II. During much of this period, rapid technological change, economies of scale and low interest and inflation rates made it possible for these companies to meet their service obligations without requiring price increases from their regulators. Regulated prices were sticky downward and little political pressure emerged for administrative reductions in prices despite the relatively high profits that these firms were able to earn under what would now be called 'fixed price' regulatory mechanisms. Although regulation was theoretically of a 'cost plus' variety, in fact there were long periods of regulatory lag that effectively operated as a de facto fixed price regulatory contract with high-powered incentives (as discussed further below). However, as economic conditions changed and nominal costs increased rapidly, political resistance to price increases intensified and triggered major changes in the nature and effects of regulation in this sector. This work suggests that not only are simple characterizations of regulation as being a pure 'cost plus' system inconsistent with

actual practice, but the nature and performance attributes of regulatory mechanisms are determined endogenously by a complex set of economic, technological and political forces.

Finally, Sam Peltzman's (1989) paper (Chapter 5, p. 123) assesses the strengths and weaknesses of Stigler's theory of regulation and related interest group politics theories in light of the experience of ten years of deregulation and regulatory reform in the US during the 1980s. This perspective does reasonably well in explaining the historical experience in some cases and less well in others. Overall, however, politics plays an important role in influencing the nature of regulation, its performance and the speed and direction of regulatory reform and deregulation.

Part II: Efficient Pricing for Natural Monopolies When Regulators are Well Informed

While political considerations play a central role in influencing the behavior and performance of regulatory agencies, it is nevertheless important to understand how economic regulatory agencies *should* establish prices assuming that they are trying to regulate in a way that maximizes the net social value of the regulated industry's activities. Regulatory actions affect the costs of providing services, the level and structure of prices that consumers pay when they seek to buy services from the regulated monopoly. From an efficiency perspective, the level and structure of prices should provide consumers with incentives to make efficient utilization (consumption) decisions regarding their use of the services provided by the regulated monopoly. Prices should also provide adequate revenues to the regulated firm to enable it to cover the (efficient) costs it must incur to provide these services. If regulatory institutions cannot credibly commit to rules that provide a reasonable opportunity to recover the costs of providing services efficiently, then private firms will be unwilling to commit capital to the regulated sectors. Credible commitments of this type are especially important when long-lived sunk investments are required to supply service economically.

Most of the initial theoretical research on pricing of natural monopoly services assumed that the goal of regulation was to establish a set of prices that would provide consumers with good price signals and that regulators were well informed about consumer demands for the relevant services and the cost of producing these services efficiently. One of the first things that students of microeconomics learn is that the efficient price is the price that equals the marginal cost of supplying output at the point where supply and demand are in balance. Applying this simple principle to many regulated sectors is challenging because of the cost attributes that we associate with natural monopoly and the unusual demand attributes which some of these sectors possess. For many infrastructure sectors, demand varies widely over the course of a day and year and supplies of the regulated services cannot be stored or rationed in a way that reflects the willingness to pay of individual consumers. As a result, demand placed by consumers on the market must clear the market based on the capacity instantaneously available to meet demand at each point in time and, when there is not enough capacity to clear demand, must be rationed more or less randomly. Accordingly, enough capacity is typically built to meet the *peak* demand expected to be made on the sector (taking into account uncertain equipment outages). Moreover, a mix of supply techniques with different capital/variable input ratios is often selected to minimize total supply costs to reflect the fact that some capacity will be operated for long durations (base load) while other capacity will be operated for short durations

(during the peak period) due to variable demands and the absence of storage. This means that the marginal cost of supplying services can vary widely with variations in demand and that both the marginal costs and the associated first-best efficient prices will vary as well over the demand cycle. Marcel Boiteux's (1960) paper (Chapter 7, p. 199), which was originally published in French in 1949 and applied to pricing of electricity by Electricité de France during the 1950s, is a classic paper that develops the theory of marginal cost pricing with variable but deterministic demands, examining cases where peak and off-peak demands are independent and cases where they are interdependent. (Closely related work was published in the US by Peter Steiner (1957) and in England by Ralph Turvey (1968).) Dennis Carlton's (1977) paper (Chapter 8, p. 222) examines cases where demand is stochastic, building on earlier work done in France by Boiteux and other researchers at Electricité de France. The many subsequent papers written on peak load pricing, and the application of these principles to tariff design in regulated industries around the world, are based on the fundamental theoretical results developed in these seminal papers.

In many cases pure marginal cost pricing does not yield adequate revenues to cover the total costs of supplying natural monopoly services, as a consequence of single or multiproduct scale economies, and the associated fact that in such cases marginal cost is below the relevant measure of the average cost incurred by the firm to provide service. As a result, pure marginal cost pricing often conflicts with budget balance or firm viability constraints. There is an extensive literature on second-best pricing that relies either on linear prices (Ramsey/Boiteux pricing) or more general non-linear pricing mechanisms (two-part tariffs and more complex schemes in which the average prices vary with the amount consumed by individual consumers). The (1970) paper by Baumol and Bradford (Chapter 9, p. 227) elaborates the theory for developing optimal second-best linear prices in the face of a binding budget constraint. The results in this paper follow directly from Frank Ramsey's (1927) paper on optimal commodity taxation and closely related work on pricing of natural monopoly services work by Marcel Boiteux published in French in 1956. These papers lead to the result that is generally referred to as 'Ramsey-Boiteux pricing'. In order to meet the firm's budget balance constraint, prices must be set above the marginal cost of supplying services. In order to minimize the loss in consumer surplus resulting from deviations of price from marginal cost to produce enough revenues to cover the firm's total costs, the Ramsey-Boiteux prices are set so that the deviations between prices and marginal costs are inversely proportional to the elasticities of demand associated with different groups of consumers and different products produced by the natural monopoly (with proper adjustment for cross-elasticities of demand when the demands are not independent). Ramsey-Boiteux pricing is effectively profit-maximizing third-degree price discrimination but where the revenues generated are capped at the firm's total supply costs. Subsequent research relaxed the assumption that the firm could only charge linear prices, allowing initially for two-part tariffs (a fixed charge for each consumer independent of usage plus a usage charge) and then for more general non-linear tariffs where the price paid for the marginal unit of consumption depends on the quantity purchased by each consumer. The (1980) paper by Mirman and Sibley (Chapter 10, p. 246) develops the theory of optimal non-linear prices for multiproduct natural monopolies. Non-linear pricing can significantly reduce the welfare losses associated with setting prices that deviate from marginal cost to meet the firm's balanced budget constraint. Brown and Sibley (1986) provide a detailed discussion of this entire literature and present a number of useful numerical examples.

Part III: Regulatory Mechanisms When Regulators are Imperfectly Informed

If the regulator had complete exogenous information about the regulated firm's production and cost *opportunities* today and in the future, about the demand structures for all types of consumers and how they will evolve over time, about the variables necessary to implement income distributional goals, and if the public were assured that the regulator could be trusted to pursue its goals fairly and efficiently, the regulator's task would be fairly straightforward. The regulator could fairly mechanically calculate the (second-best) optimal price levels and tariff structure for the regulated firm at every point in time and adjust them over time as demand and cost conditions change. Elaborate procedural safeguards would not be required because the regulator is assumed to pursue a well-defined set of 'public interest' goals efficiently.

Unfortunately, this happy situation never exists in reality, no matter how useful may be the theoretical benchmarks derived from models that assume that it does. The most fundamental problem that regulatory agencies must confront in designing regulatory mechanisms is that the regulated firm possesses more information than the regulator about its overall production cost opportunities, the costs of the individual services that it supplies, the operating characteristics of its network *ex ante* and in real time, the effort that it expends to keep costs low, the attributes of its customers' demand patterns, the quality of the services it provides and the costs of improving it, and the responsiveness of its customers to various tariff structures. That is, there is an asymmetry of information between the regulatory agency and the regulated firm. The information asymmetry can be reduced by requiring the regulated firm to report to the regulator on its costs, prices, demand patterns, technical operating characteristics, and providing for associated auditing authority. Moreover, the regulator's capability to obtain good information about cost, demand, distributional and other relevant variables may improve over time as it gains experience and its auditing and analysis capabilities grow. But the information asymmetry can never be eliminated completely. Accordingly, the regulatory game is one in which the regulated firm will always know more about its economic environment than does the regulator and will try to extract some rent from consumers as a result of its information advantage. At the same time, as the regulator endeavors to convey to consumers the benefits of lower costs it may at the same time provide incentives for the regulatory firm to produce inefficiently, dissipating some of the surplus available to consumers through wasteful expenditures. That is, the regulator must confront both adverse selection and moral hazard problems when it applies regulatory mechanisms to the regulated firm in the real world.

A great deal of the recent research on the design of 'optimal' regulatory mechanisms has proceeded within a framework where there is an asymmetry of information between the regulated firm and the regulatory agency. The goal is to develop 'incentive regulation' mechanisms that most effectively utilize the information that the regulator has available to it in a way that yields a favorable balance between the goals of extracting rents for consumers, the firm's budget balance or viability constraint, and the goal of inducing the firm to supply services as efficiently as possible given its technological opportunities and input prices. If the regulator had to be concerned only with inducing the firm to produce efficiently and assuring that the firm earned adequate revenues to cover its costs, then the regulator could set a high *ex ante* fixed price or price cap, which is not tied to the actual costs incurred by the firm and which would define the price the firm is permitted to charge (dynamically this price could change over time along with exogenous inflationary and technological parameters). In this case, the

regulated firm would be the residual claimant on any cost savings it achieves and would have high-powered incentives to minimize costs. Alternatively, if the regulator placed a lot of weight on not leaving any revenues on the table for the firm, above the firm's actual costs, the regulator could compensate the firm based on a 'cost plus' formula which would reimburse the firm for exactly (no more and no less) the costs that it incurs. The problem with the 'fixed price' regulatory contract is that it is likely to unduly burden consumers by leaving a lot of surplus to the regulated firm by allowing it to charge high prices. (Fixed price contracts can also create perverse incentives regarding the quality of service provided by a regulated monopoly if quality norms are not included.) The problem with the 'cost plus' contract is that it eradicates any incentives the firm has to produce efficiently. In a world with asymmetric information, the optimal regulatory rule is likely to be a hybrid rule that ties the payments made to the firm partly to the costs it actually incurs and partly to a fixed fee.

Regulatory practice in many countries – especially the US and Canada – during the 20th century evolved under some variant of cost-plus regulation. In the US, a form of cost-plus regulation called 'rate of return' regulation applied to legal monopolies in the electricity, telephone, natural gas pipeline and distribution and water distribution sectors. The leading theoretical paradigm that evolved for understanding the effects of rate of return regulation on the behavior of regulated firms during the 1960s and 1970s is based on the model developed by Harvey Averch and Leland Johnson (A-J) in a paper published in 1962. Many of these theoretical results as well as a number of extensions are included in the (1970) paper by Baumol and Klevorick (Chapter 11, p. 261) and in a book by Elizabeth Bailey (1973). Averch and Johnson model a neoclassical profit-maximizing monopoly subject to a constraint on the rate of return on investment that the regulated firm can earn. The regulator is extremely 'dumb and impotent' knowing absolutely nothing about the firm's cost opportunities or the demand for its services, aside from knowing the firm's cost of capital. The only regulatory instrument that the regulator has at its disposal is the ability to set the allowed rate of return at a level that must be greater than the firm's cost of capital. The primary result of this theoretical analysis is that the profit-maximizing firm is induced to substitute capital for labor inefficiently because the effective shadow price of capital faced by the firm is less than its true cost of capital as a consequence of its being subject to this rate of return constraint. Efforts to document empirically this effect have not been particularly successful (Joskow and Rose, 1989), though incentives created by rate of return regulation are likely to help to explain why firms subject to this form of regulation prefer to own facilities that produce inputs that they use to produce their primary products by vertically integrating into the production of these inputs.

The Averch-Johnson model is unsatisfactory from a number of perspectives. In practice, pure cost-plus regulation is a bit of a straw man and has probably gotten a worse reputation than it actually deserves. As applied in practice in the US, prices for regulated services are not continuously tied to accounting costs and this form of regulation is better characterized as a hybrid system that embodied both fixed price and cost-based elements (Joskow and Schmalensee, 1986 (Chapter 14, p. 338)). Depending on which point in history one examines, the period of time between regulatory reviews – regulatory lag – has often been several years. Regulatory lag effectively turns cost-plus regulation into a fixed price regulation system with a cost-based 'ratchet' adjustment taking place every few years and has similar incentive properties. Consumers left some rents on the table for the regulated firms, but the system provided reasonably good efficiency incentives and high rates of productivity growth. In addition, not only are

costs subject to elaborate accounting and auditing requirements in the US, but regulators do not have to accept the costs that are presented to them if they determine that they are 'unreasonable' or 'imprudent'. Regulators can make these judgements through management audits or benchmarking the firm's costs against comparable firms (or statistically comparable firms), effectively applying some 'yardstick competition' to evaluate the firm's costs.

The Averch-Johnson framework implicitly makes extreme assumptions about the regulator's knowledge about the firm's costs and demand opportunities and the regulatory instruments at its disposal. The A-J regulator knows nothing about the regulated firm's cost opportunities or the demand for the services that it produces, and has only the rate of return at its disposal as a regulatory instrument. Why the regulator sets the allowed rate of return at a particular level is largely outside of the model (aside from a short literature that derives the 'optimal' rate of return). As a result, the A-J model is neither a realistic positive or normative model of regulatory behavior and firm responses to it. The more recent literature on regulatory mechanism design or incentive regulation in a world where there is an asymmetry in the information available to the regulator and the regulated firm significantly improves on the A-J framework.

The papers that have developed the theory of regulatory mechanism design generally proceed under the assumption that regulators seek to pursue the public interest by maximizing a well-defined social welfare function that in one way or another gives consumer surplus more weight than producer surplus. More generally, despite differences in approach, the tradeoff between the *efficiency* with which a regulated service is produced and priced and the *distribution* of the net economic benefits of production between consumers and the regulated firm plays a central role. Baron and Meyerson's (Chapter 12, p. 290) paper (1982) is the first paper to apply the general techniques of mechanism design theory to the analysis of regulatory pricing rules. Their model assumes that the regulator cannot observe the firm's realized costs but can observe the firm's output and knows the probability distribution of the regulated firm's costs. Since cost is unobservable *ex post*, regulatory mechanisms that are based on realized costs are not available and, as a result, the optimal regulatory mechanism must take the form of a 'fixed price contract'. This in turn implies that moral hazard (production inefficiency due to distortions in managerial effort) does not exist as a potential performance problem. Relative to the full information case, in the Baron and Myerson world the regulator must set prices at levels above the second-best efficient (Ramsey-Boiteux price) level. Baron and Besanko (1984) have extended this model to allow for random audits of costs. Realized prices are lower here since prices can depend in part on observed costs and when audits are costly high cost firms are more likely to be audited than low cost firms. Lewis and Sappington (1988) (Chapter 15, p. 387) further extend this framework for situations where the firm has more information about demand than cost.

Laffont and Tirole's (1986) paper (Chapter 13, p. 310) is the first in an important series of papers that specifies and analyzes a model where the firm's cost opportunities are uncertain from the perspective of the regulator, the regulator cannot observe managerial effort, but the regulator can observe (for example via an audit) the actual cost that the firm incurs to supply output *ex post* and can tie the firm's compensation at least partially to the realized production costs. As a result, the model is set up in such a way that the regulator faces both adverse selection and moral hazard problems and the optimal regulatory contract reflects a clear tradeoff between promoting efficient production decisions by giving the firm a high-powered (fixed price) incentive scheme and extracting rents from the firm by offering the firm a cost-plus

contract. The regulator offers the regulated firm a menu of regulatory contracts that vary from a fixed-price contract to a cost-plus contract with a continuum of hybrids in between. Except at the extremes, regulated firms will generally choose a hybrid regulatory mechanism that partially ties the firm's compensation to the realized costs of production. Low cost firms are rewarded more than high cost firms reflecting the existence of both moral hazard and adverse selection problems. Laffont and Tirole's (1990) paper (Chapter 16, p. 400) extends this type of analysis to the case of multiproduct firms. This framework has been applied to a wide range of other regulatory issues including quality of service, specific investments and associated opportunism problems and even the political economy of regulatory capture. See Laffont and Tirole (1993).

The extensive theory about incentive regulation that has evolved in the last several years has met with only limited acceptance by regulators. A significant barrier to application has been the difficulty of translating the theoretical results into relatively simple implementation rules that regulators can apply in practice. Joskow and Schmalensee (1986) (Chapter 14, p. 338) point out that regulators have known and applied the basic principles of incentive regulation for over a century and discuss contemporary efforts to apply these concepts to electric utilities in the US. The most popular type of incentive regulation mechanism being used in most countries is a 'price cap' mechanism where prices are fixed at a point in time and then adjusted for several years according to a formula based primarily on exogenous variables reflecting input price changes and productivity growth norms. While the formula is in operation it is effectively a fixed price mechanism that gives the regulated firm high-powered incentives to reduce its costs. However, after a few years, price cap mechanisms typically provide for a review and these reviews typically compare prices to realized costs *ex post* and lead the regulators to reset the price level to reflect these costs. This regulatory 'ratchet' introduces a 'cost-plus' element into the regulatory mechanism that allows for a reallocation of rents from producers to consumers when realized costs are significantly below the level provided for in the price cap mechanism. It also protects the firm from contingencies in which the prices fall below the costs the firm must incur to supply output due to excessively optimistic forecasts of productivity growth, threatening the firm's continued viability. However, the ratchet necessarily softens the incentives that the firm has to devote adequate effort to cost reduction and increases the costs of moral hazard. Schmalensee (1989) (Chapter 17, p. 436) performs a simulation study that analyzes the properties of simple linear regulatory rules that include pure cost-plus and price-cap regimes as well as hybrid regimes in a world where regulators are uncertain about costs *ex ante*, can observe costs *ex post*, and (unlike in Laffont and Tirole's work) the regulator cannot make transfers to or impose taxes on the regulated firm. He finds that hybrid regulatory mechanisms in which prices depend at least in part on realized costs generally substantially outperform pure price caps without any cost-based ratchet at all. This is consistent with actual regulatory practice in many countries.

Part IV: Effects of Economic Regulation

Beginning in the 1960s, economists started to examine empirically the effects of economic regulation on prices, production costs, input prices, innovation and other performance norms. They also examined the actual behavior of regulatory agencies. This work has often played an

important role in stimulating regulatory reform, industry restructuring and deregulation. The empirical literature on the effects of regulation is voluminous. Part IV includes a set of important papers using different empirical methods and examining a variety of measures of performance. More extensive surveys can be found in Joskow and Rose (1989) and Hahn and Hird (1991).

There are several empirical approaches for measuring the effects of economic regulation on firm and industry performance. One approach is to compare the performance of firms and markets that are subject to regulation to similar firms and markets that are not subject to regulation. A closely related approach is to examine how differences in the type or intensity of regulatory instruments applied to a sector affects firm and sector performance. A third approach is to perform a 'before and after' analysis that looks at changes in various indicia of performance as firms become subject to price and entry regulation or as they are deregulated. Finally, a variety of empirical techniques are available to simulate the performance of a competitive industry with demand and cost attributes of the regulated industry being studied. The effects of regulation are then measured by comparing simulated performance to actual performance. The latter approach played an important role in making the case for deregulation of the airline industry in the US and has been used to evaluate the performance of the electric power and telephone sectors in the US.

Stigler and Friedland's (1962) study of the early 20th-century US electric power industry is a classic example of the application of the approach that compares regulated and unregulated firms (unfortunately the paper was unavailable for publication in this collection). The paper focuses on the prices charged for electricity by US electric utilities in the early decades of the 20th century, a period of time when some states had established state commissions to regulate electricity prices while others had not. Stigler and Friedland find that the prices charged for electricity by utilities in states with commission regulation are only slightly lower than the prices charged by firms in 'unregulated' states and conclude that regulation is not effective in controlling prices. However, Sam Peltzman's (1993) paper (Chapter 18, p. 459) shows that corrections of coding errors in the original Stigler/Friedland data base leads to a very different result; prices for electricity in states with commission regulation are over 20 per cent lower than in states without this type of regulation during this period. Despite these errors, the Stigler and Friedland paper is very important because it stimulated many other scholars systematically to examine the effects of economic regulation on the behavior and performance of regulated firms. By the end of World War II, almost all states had introduced commission regulation of electric utilities. Joskow's (1989) paper (Chapter 19, p. 474) follows the evolution of the effects of state commission regulation on electric utilities during the post-war period, the causes of the perceived poor performance of the industry, how performance problems created political pressure for allowing competing generating companies to enter the electricity sector, and how technological change and falling fossil fuel prices facilitated the development of a competitive electric generation sector. Similar forces are leading to the restructuring of electricity sectors around the world in ways that give unregulated competitive markets for electric generation a central role (for example Joskow, 1997).

Nancy Rose's (1987) paper (Chapter 20, p. 549) examines the effects of regulation on wages, focusing on the wages paid to drivers in the regulated trucking industry in the US. This paper integrates a before and after deregulation approach with an approach that compares regulated and unregulated firms. Rose finds that drivers in the US trucking industry were paid considerably higher wages than were drivers with similar skills and experience working in