

BONN ECON DISCUSSION PAPERS

Discussion Paper 32/2001

Regulation: Theory and Concepts

by

Dieter Bös

October 2001

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is
sponsored by the

Deutsche Post  World Net

MAIL EXPRESS LOGISTICS FINANCE

October 30, 2001

Regulation: Theory and Concepts

Dieter Bös

Department of Economics, University of Bonn, D-53113 Bonn, Germany

Introduction

Public utilities are enterprises which supply essential goods or services, where 'essential' means that they cannot be cut off without danger of total or partial collapse of an economy. From an allocative point of view these enterprises contribute to the infrastructure of the economy, while from a distributional point of view they contribute to providing consumers with necessities of life. The most important public utilities can be found in the areas of electricity, gas, water, telecommunication, postal services, radio, TV, airlines, railroads and urban public transportation. It is not the ownership but the lack of competition which justifies regulation of the activities of public utilities. Accordingly, privatization does not necessarily imply the end of government regulation. If it is impossible to expose a public utility to competition, then price and quality regulation typically are regarded as inevitable, in spite of the government's interest in withdrawing from intervention in the particular field as signalled by the very act of privatization.

This raises the question of how far competition can be introduced in the supplies of telecommunication, rail, and the like. For a long time this question was not asked, because all public utilities were thought to be 'natural monopolies,' characterized by a subadditive cost function¹ and by sustainability:²

¹Good overviews over the precise meaning of subadditive cost functions can be found in Panzar (1989), pp. 23-33, or Sharkey (1982), pp. 54-83.

²See, in particular, Baumol et al. (1982). A very clear treatment of the problem of the contestability of monopolies can be found in Sharkey (1982), chapter 5.

it is cheaper to produce goods by a monopoly than by many firms, and potential market entrants can be held off without predatory measures. In such cases unregulated private enterprises would exploit the market. Therefore, regulation is necessary. Contestability of monopolies is particularly relevant in the case of multi-product enterprises where market entry may refer to only one product or a subgroup of products of the public utility. Obviously, only rarely does the whole public utility exhibit the properties of a natural monopoly. In many cases only a network is a natural monopoly and competition is possible with respect to the other parts of production or distribution of services. Various companies which administrate rolling stock could use the same railroad infrastructure. Various telecommunication companies can use the same network. Typically, however, a relatively long time is needed to successfully establish competition in these fields. The British market for telecommunications was characterized for a long time by British Telecom as the market leader and Mercury as the follower. Note that telecommunications is a good example for an erosion of natural-monopoly positions: the invention of the mobile phone practically implies that for particular services no network is needed. This change in the technology, accordingly, leads to fierce competition in this field. A failure story, on the other hand, is the privatization of British Rail. Here, the railroad network was considered a natural monopoly that had to be regulated, but was not regulated carefully enough and caused quality to deteriorate quickly. The train-operating companies responsible for providing passenger services have route monopolies under franchises.

Whenever there is no competition, or this competition is not strong enough to prevent a public utility from exploiting its customers, government regulation of public utilities remains on the agenda of economic policy and, therefore, remains an important subject of economic theory.

A very general framework for a theory of regulation of a public utility is as follows. The theory considers a two-person game between a regulator and a manager who represents the public utility. The objectives of the two players may vary. The regulator may alternatively be considered as a welfare maximizer, as a politician who wants to maximize votes, or as a bureau-

crat who wants to maximize his power and, therefore, is interested in a high budget of the utility or in high output. The manager may be modelled as a profit maximizer or as an agent who is interested in his personal income and in the disutility of his effort. In this two-person relationship the players have to consider various constraints. First, the utility is typically obliged to meet all demand at the regulated prices.³ Second, in many cases the public utility is explicitly restricted in its ability to maximize profits. Alternative regulatory models deal with direct profit constraints (profit must be smaller than some exogenously given threshold), or with indirect profit constraints, imposing ‘caps’ on prices or an upper boundary on the rate of return on investment. Third, the regulation must not eliminate the public utility. Regulation has failed if the regulator drives the utility into bankruptcy or, in an alternative formulation, induces the manager to leave his job (violation of the manager’s participation constraint). Finally, whenever the regulator is not fully informed, he must induce the manager to operate in line with the regulator’s intentions (incentive compatibility).

Any regulatory activity should be evaluated by comparison with a benchmark model. A fully informed welfare-maximizing regulator may be taken as the basis of such a benchmark. Which prices would he impose on the public utility? Let me present two very simple rules; many more refined rules can be found in Bös (1994). Consider first the simple maximization of the sum of consumer and producer surplus,⁴ without any profit constraint. In this case the regulator will choose prices which are equal to the respective marginal costs. If the public utility operates under increasing returns to scale, these marginal-cost prices will lead to a deficit of the firm. If this deficit is considered too high, the regulator may choose prices which maximize the sum of consumer and producer surplus under a revenue-cost constraint.⁵ In this

³In the peak-load pricing literature it has been shown that under particular assumptions rationing of the demand may be welfare-optimal. See Bös (1994), chapter 15.

⁴The terms producer surplus and profit are used synonymously in this paper.

⁵This constraint requires either that the deficit should not be too high, or that the firm’s revenue should at least cover the production costs, or that a minimum profit should be attained.

case, Ramsey prices will be chosen.⁶ Consider a two-product utility,⁷ selling quantities x_1 and x_2 at prices p_1 and p_2 . Denote marginal costs by $C_i, i = 1, 2$. Ramsey pricing is characterized by the following marginal conditions:

$$\frac{p_1 - C_1}{p_1} = -\lambda \frac{\epsilon_{22} - \epsilon_{12}}{\epsilon_{11}\epsilon_{22} - \epsilon_{12}\epsilon_{21}}; \quad \frac{p_2 - C_2}{p_2} = -\lambda \frac{\epsilon_{11} - \epsilon_{21}}{\epsilon_{11}\epsilon_{22} - \epsilon_{12}\epsilon_{21}}, \quad (R)$$

where $\epsilon_{ij}, i, j = 1, 2$ are price elasticities of demand⁸ and $\lambda \in [0, 1]$ is a scale parameter which depends on the profit threshold. If $\lambda = 0$, we have marginal-cost prices, if $\lambda = 1$, we have monopoly prices. Now consider the right-hand side of the first of the above equations

$$-\lambda \frac{\epsilon_{22} - \epsilon_{12}}{\epsilon_{11}\epsilon_{22} - \epsilon_{12}\epsilon_{21}} = -\frac{\epsilon_{22}/\lambda - \epsilon_{12}/\lambda}{\epsilon_{11}/\lambda \cdot \epsilon_{22}/\lambda - \epsilon_{12}/\lambda \cdot \epsilon_{21}/\lambda} = -\frac{\eta_{22} - \eta_{12}}{\eta_{11}\eta_{22} - \eta_{12}\eta_{21}},$$

with $\eta_{ij} := \epsilon_{ij}/\lambda > \epsilon_{ij}$. A similar transformation holds for the second equation. Accordingly, the Ramsey utility behaves like a profit-maximizing monopolist who overestimates all price elasticities of demand by the same factor $1/\lambda > 1$.

Overestimation of elasticities implies that the utility will be more cautious than the profit-maximizing monopolist when it comes to raising prices above marginal costs: the prices are set more cautiously, the more easily demand is lost in the case of a price increase (and this is just the same problem a pure monopoly faces, hence the possibility to characterize Ramsey prices by a comparison with monopoly prices). In contrast to monopoly prices, however, the cautious behavior of the Ramsey firm implies a lower price level than for a profit-maximizing monopolist, resulting from the fact that the profit constraint is lower than the monopoly profit (and higher than the deficit at

⁶Ramsey (1927) considered the problem where a given tax revenue should be raised by indirect taxation at minimal welfare loss. However, for given producer prices indirect taxation means a choice of consumer prices, and, accordingly, his theory can directly be transferred to the case of regulation of the prices of public utilities. It was Boiteux (1956, 1971) who first presented a general-equilibrium model on public-sector pricing with a given profit constraint.

⁷The extension to the n -good case is straightforward. See Bös (1994), chapter 8.

⁸We deal with compensated price elasticities, that is, the elasticities are defined along Hicksian demand functions $x_i^h = x_i^h(p_1, \dots, p_n, u^h)$, where u^h is the utility of the h -th consumer, $h = 1, \dots, H$.

marginal-cost prices). The most popular special case of Ramsey pricing is the so-called ‘inverse-elasticity rule:’ if all cross price elasticities are ignored, then the relative deviation of any price from the marginal costs is lower, the higher the direct price elasticity of demand.

In contrast to the benchmark models, regulators are in practice never fully informed. There is, first, the moral-hazard problem, which arises if the regulator and the manager of the public utility are equally badly informed when the decision on regulation is made. In this case both allocative and productive efficiency can be achieved by ‘selling the store to the agent’ and stipulating a Loeb-Magat (1979) mechanism: the regulator gives the consumer surplus to the manager in exchange for a lump-sum compensation. The manager is allowed to retain the profit. Therefore, he will maximize the sum of consumer and producer surplus and attain the first best. In practice, this mechanism is not applied, because it shifts all of the risk to the manager and away from the regulator, and because it is too expensive in the case of asymmetric information, which is more plausible than symmetric information: it is highly likely that the manager of the public utility is better informed than the regulator. Therefore, most of the modern theory of regulation concentrates on the adverse-selection problem, where the information is asymmetrically distributed when the decision on regulation is made. It has been shown that there is a special class of contracts between the regulator and the public utility whose result is always at least as good for the regulator as any other contract he could conceive. In this contract the manager is asked to announce the actual value of his private information and gets a specially designed incentive income which induces truthful revelation (‘revelation principle’). The incentive income implies that the manager of the public utility gets an information rent for revealing his private knowledge. Seminal work on price regulation under asymmetric information is due to Baron and Myerson (1982) and Laffont and Tirole (1993).

This article is organized as follows: we begin with the treatment of price regulation by simple regulatory rules. These rules require a minimum of information on the side of the regulator. In particular, he need not know the functional shapes of demand and cost functions or the probability distri-

bution of some unobservable variable. Unfortunately, the simple regulatory rules typically raise negative incentive effects on the side of the regulated utility and, therefore, they are only rarely applied in economic practice. We shall discuss the iterative mechanisms of Vogelsang and Finsinger and the yardstick regulation. Then we turn to informationally-demanding price regulation, which avoids the negative incentive effects, but requires much more information on the side of the regulator and, therefore, is more of a fascinating theoretical exercise than an actually applied instrument of economic policy. Therefore, we devote another section to price-cap regulation, which is practically applied, obviously because it implies a satisfactory compromise between information requirements for the regulator and negative incentive effects for the public utility. Finally, we turn to some problems of quality regulation. A brief conclusion follows.

Simple regulatory rules

Regulation by an iterative process

Consider the following regulatory adjustment process which leads to Ramsey prices.⁹ Players of the game are a profit-maximizing public utility under increasing returns to scale, and a welfare-maximizing regulator who has only minimal information about the activities of the utility. At the beginning of a period the regulator stipulates a set of prices which are at most cost-covering if applied to the quantities sold in the period before. Within this set the utility chooses those prices which maximize its profit; this profit may well be positive. The profit-maximizing prices and quantities of the present period serve as the basis for the regulatory set of prices of the next period, where the utility once again chooses profit-maximizing prices that belong to the regulatory set. This iterative process continues until break-even Ramsey prices are achieved.

Why can such an iterative process lead to optimal prices? Recall the Ramsey benchmark model. The optimal prices resulted from a maximization

⁹See Vogelsang and Finsinger (1979).

of the sum of consumer and producer surplus for a given profit constraint. By duality, the same prices result if the profit is maximized under the constraint that consumer surplus plus producer surplus should not fall below an adequately chosen threshold. Now consider the Vogelsang-Finsinger model. The utility maximizes profit. This maximization is constrained by a regulatory set of prices which are at most cost-covering if applied to the sales of the previous period. It is obvious that this regulation protects the consumers against exploitation from the profit-maximizing utility. Therefore, it has the same function as the minimum threshold on consumer plus producer surplus.¹⁰

The main advantage of this regulatory adjustment process is the minimal information requirement for the regulator. In order to stipulate the regulatory set of prices, he only has to know the prices, the quantities and the total costs of the past period. In particular, he does not need any information about the total shape of demand and cost functions, and he does not need any information about the distribution of particular non-observable variables. On the other hand, there are various disadvantages of the regulatory adjustment process. The utility may have an incentive to increase costs in the long-run because waste today leads to a higher price level tomorrow and increases the long-run profits.¹¹ Moreover, the demand and cost functions must remain unchanged until the Ramsey optimum is achieved and this can only hold if the revision of the regulatory set of prices is made fairly frequently. Even annual revisions may be too infrequent.

Yardstick regulation

Yardstick regulation can be applied by a regulator who faces various similar utilities and, therefore, can use the information about one utility to regulate the others. The regulator asks one utility for the actual value of some variable which is private knowledge of this utility. However, he commits himself

¹⁰It can be shown that the regulatory set of prices is tangent to the indifference surface of the welfare function. The convexity of welfare allows substitution of this tangent hyperplane for the actual welfare function in the various steps of the iterative process.

¹¹See Sappington (1980).

to use this piece of information only for the regulation of all the other utilities, not for the regulation of the particular utility itself. This utility, in turn, is regulated on the basis of the information acquired from all the other utilities. Since the utility knows that telling the truth will not influence its own regulation, it has no incentive to give false information. Hence, it will tell the truth. Yardstick regulation applies a basic idea which has been used often in mechanism-design literature, in particular in the mechanisms for the revelation of preferences for public goods.

A more detailed analysis of the yardstick mechanism is as follows.¹² Assume that there are n identical regional monopolies. The demand function is the same in every single region. The firms operate under constant production costs. However, these production costs can be reduced by R&D investments. A welfare-maximizing regulator sets the prices and a subsidy which is paid to each utility. If the regulator were fully informed, he would choose marginal-cost prices and equate the subsidy to the costs of the R&D investments. However, the regulator does not know the R&D technology. Therefore, he applies the following mechanism:

- At date 1 he announces the regulatory rules: for any single utility he will set prices that are equal to the mean of all other utilities' announcements of production costs. Every utility will receive a subsidy which is equal to the mean of all other utilities' announcements of the R&D investment costs. He also commits not to bail out any utility in the case of bankruptcy.
- At date 2 each firm invests in R&D and the regulator comes to know their investment costs and the associated production costs. This is made possible because under the announced regulation no firm has an incentive to hide information on R&D or production costs.
- Given his information on investment and production costs, at date 3 the regulator actually fixes the price and the subsidy for any single firm according to the regulatory rules announced at date 1.
- Finally, at date 4 the utilities produce, sell their products at the regulated prices and encash the subsidies.

¹²The standard paper on this mechanism is Shleifer (1985).

This mechanism implements the first best, that is, marginal-cost prices and subsidies which cover the costs of the R&D investments.

The main advantage of this mechanism is its low information requirement for the regulator. He does not need any information about cost and demand functions. He just applies the insight that no firm is interested in cheating unless this improves the profit. Since the own announcement has no influence on the profit, no firm will cheat and the regulator gets all the information he needs for a first-best regulation. Moreover, the achievement of the first best is driven by the firms' profit-maximizing behavior and, therefore, there are no adverse incentive effects which might stop the firms from choosing their strategies which lead to the first best.

Unfortunately, however, yardstick regulation also has quite a few disadvantages. First, it is vulnerable to collusion, because collusion makes profits dependent on own announcements. This makes yardstick regulation questionable in all those cases where various privatized utilities have been created by splitting up the former monolithic publicly owned utility.¹³ Similarly, for effective yardstick competition, there must be a number of firms in the industry with similar demand and cost conditions. This is the reason the UK regulators have opposed some proposed mergers in the electricity and water sectors. Second, it is difficult to understand how a regulator of a privatized utility can commit himself not to bail out a utility which he has driven into bankruptcy by his regulatory policy. (Regulators often have a legal requirement to ensure that the regulated firm can earn sufficient revenues to carry out its proper functions.) Third, the whole merits of using this form of regulation are called into question at the practical level if cost and demand functions are different.¹⁴ This has drawn UK regulators into heated arguments with companies about the value of comparative competition.

¹³A good example are the British regional electricity companies.

¹⁴For this case Shleifer (1985), pp. 324-325, suggests a reduced-form regulation which uses predicted costs on the basis of a regression analysis linking marginal costs and exogenous characteristics of all utilities. However, the first best will then only be achieved if the regression explains 100 per cent of the variance of costs, which typically will not be the case.

Informationally-demanding regulatory rules

The principal-agent model

The regulator as principal of the game is not able to produce the firm's outputs, so he needs the manager of the public utility as his agent. There is asymmetric information. Only the manager knows the actual realization of a one-dimensional characteristic θ which influences the costs or the demand.¹⁵ We normalize θ by defining $\theta \in [\underline{\theta}, \bar{\theta}]$, where $\underline{\theta}$ is the worst case. Asymmetric information does also prevail with respect to the manager's effort: it cannot be observed by the regulator. However, the above assumptions do not imply that the regulator is ignorant of the utility's special features. Far from it! He is assumed to be very well informed. This is the serious weakness of the informationally-demanding regulatory rules compared with the simple rules of the preceding section. The regulator has to know the functional shapes of the public utility's cost and demand functions and of the manager's utility function. Moreover, he has to know the distribution function of the unobserved characteristic that influences costs or demand. Finally, it is assumed that the regulator ex post observes total costs¹⁶ or at least the produced quantities.¹⁷ The regulator's lack of information, therefore, refers only to the actual realizations of the managerial effort and the cost or demand characteristic. However, this very lack of information prevents the regulator from calculating in how much total costs or total sales result from the agent's effort or from the actual realization of a cost or demand characteristic. Therefore, the agent can cheat. The agent's utility is $U(t, e)$, where t is the managerial income and e is the managerial effort, $U_1 > 0, U_2 < 0$: the agent feels better if he gets a higher income and if he expends less effort. Therefore, the manager has an incentive to pretend that there have been adverse cost

¹⁵Of course such a characteristic could also refer to other functions which are relevant for the utility. By way of an example, in Bös (1994), chapter 31, a model is presented where such a characteristic refers to a budget-appropriation function.

¹⁶This assumption is typical for the Laffont-Tirole (1993) approach.

¹⁷Baron and Myerson (1982) wrote about regulation with unknown costs. However, they assume that the quantities are ex-post observable. For a nice presentation of the Baron-Myerson model see Laffont and Tirole (1993), pp. 155-158.

or demand shocks and, therefore, his effort had to be very high so that he should be compensated by a much higher income. What should the regulator do in such a situation? The principal-agent theory proposes the following sequence of strategic moves.

- Stage 1: The manager is better informed. Only he knows the cost or demand characteristic. The regulator only knows the distribution function of this characteristic.
- Stage 2: The regulator offers a contract which implements a direct mechanism: the manager will have to announce the actual realization of the unobservable characteristic. For every possible announcement $\hat{\theta}$ the contract stipulates an incentive income $t(\hat{\theta})$ which is defined so as to fulfill two requirements. First, the contract is *incentive compatible*, that is, the manager achieves highest personal utility if he truthfully informs the regulator, $\hat{\theta} = \theta$. The incentive-compatibility condition requires that the managerial utility is strictly increasing in the characteristic θ , that is, $U_\theta > 0$: when the manager is asked for the correct value of θ , he must not have an incentive to cheat by announcing a lower θ than actually realized. Second, the contract takes care of the manager's *participation constraint*; managerial income and effort are traded-off in such a way that it is attractive for the manager to stay at his job and not to leave to an outside position. The managerial utility $U(t, e)$ has to exceed the reservation utility \bar{U} which is the highest utility level the manager could earn at an alternative job. In a full-information benchmark the participation constraint is always binding:

$$\bar{U} = U(t^*(\theta), e^*(\theta)), \quad \forall \theta \in [\underline{\theta}, \bar{\theta}],$$

where $e(\theta)$ means that the regulator correctly anticipates how the manager will adjust his effort to the actual realization of θ ($= \hat{\theta}$). In the case of asymmetric information, however, the participation constraint binds only at the worst situation:

$$\bar{U} = U(t(\underline{\theta}), e(\underline{\theta})) < U(t(\theta), e(\theta)), \quad \forall \theta > \underline{\theta}.$$

This result is rooted in the incentive-compatibility constraint which requires utility to increase in θ . Hence, the participation constraint can only bind at

the lowest realization of θ .

It is a further part of the contract that the income will be paid at the end of the game, but only if the produced outputs are exactly equal to those quantities which the regulator has calculated on the basis of the truthful information from the manager.¹⁸ This calculation also allows the regulator to announce the prices at which the outputs are to be sold.

- Stage 3: The manager informs the regulator about the actual realization of the cost or demand characteristic.
- Stage 4: The manager chooses his effort depending on the actual value of the characteristic, $e(\theta)$.
- Stage 5: The manager produces and sells the products at the regulated prices. He encashes his income.

Asymmetric information on costs¹⁹

Assume the following cost function:

$$C = C(x_1, \dots, x_n, e, \theta); \quad C_i := \partial C / \partial x_i > 0, \quad \partial C / \partial e < 0, \quad \partial C / \partial \theta < 0.$$

Total costs depend on the vector of produced quantities x_1, \dots, x_n , the managerial effort e and an exogenous cost characteristic θ . This characteristic refers to the type of utility, from high-cost firms to low-cost firms: a particular set of output quantities requires high costs if θ is low, but low costs if θ is high. Now consider a regulator who maximizes welfare and takes account of market-clearing conditions and of a profit constraint. Furthermore, he writes a contract with the manager which is incentive compatible and fulfills the manager's participation constraint. It can be shown that in this case the regulator chooses a special type of Ramsey prices. As in the full-information benchmark model, he operates like a profit-maximizing monopolist who overestimates all price elasticities by the same factor. In

¹⁸This assumes a modelling where the regulator ex post observes produced quantities. If he ex post observes the realized costs, but not the individual quantities produced, a similar story can be told.

¹⁹This subsection presents only a very rough sketch of the relevant problems. For details see Bös (1994), chapters 28 and 29.

the special ‘inverse-elasticity’-case, as in the benchmark, the relative deviation of any price from the marginal costs is lower, the higher the direct price elasticity of demand. However, the marginal costs in the asymmetric-information Ramsey formula comprise both the marginal production costs and an incentive-correction term which copes with the manager’s incentive-compatibility problem.

The asymmetric-information Ramsey formula differs from the benchmark formula (R) by the inclusion of incentive-correction terms I_i , $i = 1, 2$. It runs as follows:²⁰

$$\frac{p_1 - C_1 - I_1}{p_1} = -\lambda \frac{\epsilon_{22} - \epsilon_{12}}{\epsilon_{11}\epsilon_{22} - \epsilon_{12}\epsilon_{21}}; \quad \frac{p_2 - C_2 - I_2}{p_2} = -\lambda \frac{\epsilon_{11} - \epsilon_{21}}{\epsilon_{11}\epsilon_{22} - \epsilon_{12}\epsilon_{21}}.$$

The incentive-correction terms I_i result from the differentiation of the manager’s incentive-compatibility constraint with respect to the i ’th quantity. Therefore, instead of considering the marginal production costs C_i , in the asymmetric-information setting the regulator considers modified marginal costs $C_i^M = C_i + I_i$.

How should one interpret the regulator’s pricing decision? Will asymmetric-information Ramsey prices be higher or lower than Ramsey prices in a full-information benchmark (assuming identical profit constraints)? A first guess would hint at higher prices, because the badly informed regulator has to pay for the production costs plus the information rent of the manager. The fully-informed regulator does not pay such a rent. It would be plausible to assume that in the case of asymmetric information the manager always gets a higher income which would enforce higher prices. Typically, however, this simple plausibility is incorrect.²¹ Since the incentive-compatibility condition requires $U_\theta > 0$, the managerial income at some level θ^o influences all incomes at higher levels of θ . This external effect is present for all realizations of θ except the best one, $\bar{\theta}$. Hence, at this point the income is chosen by the regulator so as to attain full efficiency and, therefore, the effort level is

²⁰Once again, the extension to the n -good case is straightforward, see Bös (1994), pp. 316-320.

²¹For the following treatment of managerial incomes see Bös and Peters (1991), pp. 39-41.

$$e(\bar{\theta}) = e^*.$$

Recall that the manager's participation constraint is not binding at $\bar{\theta}$ (in contrast to the full-information benchmark),

$$U(t(\bar{\theta}), e^*(\bar{\theta})) > U(t^*(\bar{\theta}), e^*(\bar{\theta})) = \bar{U},$$

where in the first term we have substituted $e = e^*$. However, this implies

$$t(\bar{\theta}) > t^*(\bar{\theta}).$$

In a low-cost firm the managerial income will be higher than the benchmark income.

Consider next the worst possible case, $\underline{\theta}$. Here it would be too costly for the regulator to enforce efficient effort (because of the external effect on all other incomes). Therefore, he settles for an effort lower than efficient,

$$e(\underline{\theta}) < e^*.$$

The participation constraint is binding,

$$U(t(\underline{\theta}), e(\underline{\theta})) = U(t^*(\underline{\theta}), e^*(\underline{\theta})) = \bar{U}$$

and, therefore, we have

$$t(\underline{\theta}) < t^*(\underline{\theta}).$$

In a high-cost firm the managerial income will be lower than the benchmark income. Therefore, the managerial income can be lower or higher than the benchmark income. Consequently, asymmetric information can imply a lower or a higher average of prices depending on whether we have a low-cost or a high-cost firm.

However, a lower average of prices does not necessarily imply that all prices must be lower than their benchmark equivalents. A particular price will be higher than its benchmark equivalent if the marginal rate of transformation between the managerial effort and the cost characteristic responds

positively to an increase in the supply of the respective good.²² An increase in output in this case makes it easier for the manager to transform exogenous costs shocks into rents.

The incentive-correction term which increases the marginal costs in the modified Ramsey rule may vanish in particular cases. First, consider the manager's trade-off between effort and the cost characteristic. This trade-off indicates how far the manager can reduce his effort if the cost characteristic is improved. The incentive-correction term vanishes if this trade-off does not depend on the supplied quantities.²³ In this case prices determine quantities, but not the relationship between effort and the cost characteristic. Therefore, there is no incentive-correction term in the Ramsey formulas. Second, consider the best realization of the cost characteristic. Here, the regulator can choose the efficient solution since there is no external effect on larger values of θ . Therefore, in the case of $\bar{\theta}$, there is no incentive-correction term in the Ramsey formula. Note that identical Ramsey formulas for regulatory prices typically will not imply identical prices in the benchmark and in the asymmetric-information case. The cost characteristic will continue to influence the managerial effort and income, and the managerial income enters the profit constraint that determines the revenue that must be raised at the regulated prices. Hence, the absolute values of the prices will be influenced by the cost characteristic even though the Ramsey structure of prices is the same in the benchmark and in the case of asymmetric information.

Asymmetric information on demand²⁴

Assume the following compensated demand functions:

$$x_i^h = x_i^h(p, u^h, e_i, \theta), \quad \partial x_i^h / \partial e_i > 0, \partial x_i^h / \partial \theta > 0.$$

²²For constant total costs and constant quantities we consider the total differential of the cost function $C(x_1, \dots, x_n, e, \theta)$. We obtain $de/d\theta = -C_\theta/C_e = -MRT(e, \theta)$ and, therefore, $MRT > 0$. If $\partial MRT/\partial x_i$ is positive, then the incentive-correction term I_i is positive, which implies a tendency towards a higher price of good i .

²³This is the case if the cost function is $C(x_1, \dots, x_n, f(e, \theta))$.

²⁴Once again this subsection presents only a very rough sketch of the relevant problems. For details see Bös (1994), chapters 28 and 30.

The quantity of good i which individual h buys depends on the vector of consumer prices p , on his utility u^h , on a demand characteristic θ and on the marketing efforts which the manager of the public utility devotes to good i , that is, e_i . The derivatives with respect to prices and utility are assumed to follow the usual microeconomic convention.

Once again, the welfare-maximizing regulator will choose a modified Ramsey rule. However, the most plausible modification does not hold. Since the manager gets an information rent, one would have assumed that, once again, the regulator considers modified costs consisting of production costs plus costs of ‘buying’ the information on the demand characteristic. This is not the case, however. The modification occurs at the demand side: the factor by which the elasticities are overestimated in the Ramsey formula is changed by incentive-correction terms. For the two-good case the modified Ramsey formula is as follows:²⁵

$$\frac{p_1 - C_1}{p_1} = - \frac{(\lambda - I_1)\epsilon_{22} - (\lambda - I_2)\epsilon_{12}}{\epsilon_{11}\epsilon_{22} - \epsilon_{12}\epsilon_{21}}, \quad \frac{p_2 - C_2}{p_2} = - \frac{(\lambda - I_2)\epsilon_{11} - (\lambda - I_1)\epsilon_{21}}{\epsilon_{11}\epsilon_{22} - \epsilon_{12}\epsilon_{21}}.$$

There are different incentive-correction terms for the different goods, so the extent of overestimation differs depending on how price changes influence the manager’s marginal disutility of effort and his trade-off between effort and the demand characteristic (how far he can reduce his effort if the demand characteristic improves).

In contrast to the case of asymmetric information on costs, decreasing marginal costs may require a totally different regulatory policy than that described in the preceding paragraph. In a simplified example²⁶ the incentive compatibility of the regulatory scheme requires prices which are increasing in

²⁵The extension to the n -good case is straightforward, see Bös (1994), pp. 336-339.

²⁶See Lewis and Sappington (1988); for a particularly simple presentation of the problem see Bös (1994), pp. 303-304. Formally, consider the second-order condition of the managerial revelation problem. The managerial utility U depends, inter alia, on the announced value of the demand characteristic, called $\hat{\theta}$. To make truthful revelation a managerial-utility maximum, we must have $U_{\hat{\theta}} = 0$ at $\hat{\theta} = \theta$, where θ is the actual value of the demand characteristic. Furthermore, we must have $U_{\hat{\theta}\hat{\theta}} < 0$. It is comparatively simple to find plausible assumptions for this second-order condition to hold in the case of asymmetric information on costs. However, this is not the case if there is asymmetric information on demand. Compare Bös (1994), p. 311 (cost side) and pp. 331-332 (demand side).

demand. On the other hand, first-best marginal-cost prices would have to be decreasing in demand because of the decreasing marginal costs. Accordingly, from the welfare point of view incentive-compatible price regulation becomes too costly. It can be shown that in this case it is optimal for the regulator to implement the same price for all realizations of the demand characteristic ('bunching'). It is too costly to elicit the true information from the manager of the firm and the price regulation is only based on the regulator's imperfect information.

Price-cap regulation

The most widely used form of price-cap regulation is the $RPI - X$ formula: a price index of the monopolistically supplied goods of a public utility must not increase by more than the retail price index minus a constant X which has been set by the regulator.²⁷ The constant X was conceptualized as a factor that measures productivity increases of the public utility. These increases should be passed on to the consumers. The productivity increases may refer to an outward-shifting production frontier that is due to technical progress. Accordingly, telecommunications should have a high X , gas should have a low X . Productivity increases may also reflect that the firm has reduced slack in its production, producing nearer to the frontier than before (approaching productive efficiency).²⁸ This argument was often put forward when price caps were introduced in the course of privatization. The constant X should also take account of demand increases that allow price reductions in the case of increasing returns to scale.²⁹ In the regulatory practice, however, several other criteria have influenced the choice of X .³⁰

²⁷This form of regulation has been proposed by Littlechild (1983).

²⁸Leibenstein (1966) coined the term *X-inefficiency* for production below the frontier. Note that the X in *X-inefficiency* has nothing to do whatsoever with the X in the $RPI - X$ formula. To avoid misunderstanding, as a synonym for *X-efficiency* we will use the term *productive efficiency* in the text.

²⁹See Vickers and Yarrow (1988), pp. 214-216.

³⁰Many further problems in the practical application of the $RPI - X$ formula are treated in Bös (1991), pp. 67-68.

- (i) Regulators often choose X so as to determine the profits of the public utility: X is increased if the profits have been high. If in such a case the regulator sets X so as to allow a fair rate of return to the firm, then the $RPI - X$ regulation comes close to a rate-of-return regulation.
- (ii) If $RPI - X$ is introduced on the occasion of privatization of a public utility, the government will have an incentive to choose a low X because this increases the profits of the utility and, therefore, the revenue that the government gets from the sale of the shares of the firm.
- (iii) How X is set depends decisively on the informational status of the regulator. The worse he is informed about costs or demand, the lower the X he must choose. Otherwise, the regulator could drive the firm into bankruptcy. At high levels of uncertainty, cost-plus regulation may be preferable to price-cap regulation, since in such a case price-cap regulation implies the concession of higher prices than cost-plus regulation.³¹

The constant X will be reviewed at regular intervals to cope with changes in the profitability of the public utility (*regulatory lag*). This lag implies a tension between achieving (and maintaining) allocative efficiency and the attainment of productive efficiency. Lags in adjusting price caps give the profit-maximizing public utility incentives to improve productive efficiency but at the cost of allocative efficiency. Consider a regulator who has chosen a particular value of X and a regulated utility which reduces its costs to increase the profit. The firm may retain this higher profit. However, at the next revision of X the prices are set so as to shift the gains from the efficiency increases from the producer to the consumers. The firm's incentives depend on the length of the regulatory lag. If the interval between two revisions is too short, there will not be many incentives for innovative activities of the utility. If the interval is too long, too much profit goes to the firm and the consumers are exploited.

If the regulator is imperfectly informed about the costs, the firm will make strategic use of the regulatory lag. Let us assume that the manager of the firm knows that the regulator will choose X so as to siphon off the utility's

³¹See Schmalensee (1989).

profits. Insofar as cost-reducing innovations are reversible, the firm has an incentive to be a high-cost firm at the moment of the regulatory reviews, but a low-cost firm in between. A sawtooth profile of the firm's cost-reducing innovations will result.³² The issue of the timing of productive-efficiency gains and price-cap reviews can be overcome (to a degree) by the use of 'glidepath' and similarly lagged adjustments of the cap rather than loading all of the adjustment into current price.

On the side of the firm, there may be imperfect information about the date of the next revision. If the manager of the utility knows some exogenous probability of the regulatory revision, he will act too cautiously in his innovative policy. A better result is achieved if the probability of revision is endogenized. This is the case if the manager knows that a revision becomes highly likely if the profit exceeds a particular level that is considered fair by the regulator. The regulation converges to prices where there is no excess of current over fair profit. Moreover, cost minimization is achieved.³³

Quality regulation

There have been many complaints about quality deterioration due to privatization and insufficient quality regulation. The UK rail privatization provides the most recent example. Quality regulation is more complicated than price regulation, because quality typically is multidimensional in nature. By way of example, the quality of local transportation services should be measured by reference to the per cent of cancelled trains, waiting time (frequency of services), travelling time, comfort of rolling stock, and cleanliness of the stations. The multidimensionality makes it impossible to find simple regulatory rules for quality regulation (like $RPI - X$ for price regulation). Simple rules can only be found if one-dimensional quality indicators are considered, for instance the reliability of supply measured in per cent of cancelled trains, or in per cent of breakdowns of electricity supply. Multidimensionality, however, implies weighting of various quality indicators, which in practice is a

³²See Armstrong et al. (1991).

³³This has been proved for rate-of-return regulation by Bawa and Sibley (1980).

complicated cost-benefit analytical task.

From the theoretical point of view, the conventional neoclassical models could well be augmented in order to deal with both price and multi-dimensional quality regulation. However, to simplify the treatment in this paper we only consider one-dimensional quality indicators q_i which enter in the cost function C , the compensated demand functions x_i^h , and the consumers' expenditure functions r^h .³⁴ For a two-good firm we have the following specifications:³⁵

$$\begin{aligned} C &= C(x_1, x_2, q_1, q_2); & C_i &:= \partial C / \partial x_i; & C_{q_i} &:= \partial C / \partial q_i, & i &= 1, 2, \\ x_i^h &= x_i^h(p_1, p_2, q_1, q_2, u^h); & \phi_{ij} &:= \partial x_i / \partial q_j, & x_i &:= \Sigma_h x_i^h, & i &= 1, 2, \\ r^h &= r^h(p_1, p_2, q_1, q_2, u^h); & Q_i &:= (1 - \lambda) \Sigma_h \partial r^h / \partial q_i < 0, & i &= 1, 2. \end{aligned}$$

Let us begin with two full-information benchmark models where quality regulation adjusts to marginal-cost pricing and to Ramsey pricing. In the first case every single quality has to be expanded until the marginal quality costs equal the sum of marginal utility gains as measured by changes of the individual expenditure functions,

$$C_{q_i} = - \Sigma_h \partial r^h / \partial q_i, \quad i = 1, 2.$$

Accordingly, the individual marginal utility gain can be interpreted as a marginal rate of substitution between quality and the individual income, where the income is measured by the expenditure function. The first-best qualities, therefore, require the equality of a marginal rate of transformation (marginal quality costs) and the sum of individual marginal rates of substitution. This condition resembles the Samuelson condition on public goods. In the second case the qualities are adjusted to Ramsey prices. The best interpretation of the optimal qualities can once again be given by a comparison with a monopolist who chooses prices and qualities so as to maximize

³⁴This presentation follows Bös (1994), chapter 16. For an explicit treatment of quality regulation see also Laffont and Tirole (1993), chapter 4.

³⁵To further simplify the formal analysis, we suppress the dependence of x_i^h and r^h on the prices and qualities of goods other than the two goods produced by the public utility in question.

his profits. The respective first-order condition on quality of good 1 is as follows:

$$p_1 - C_1 = \frac{(C_{q1} + Q_1)\phi_{22} - (C_{q2} + Q_2)\phi_{21}}{\phi_{11}\phi_{22} - \phi_{12}\phi_{21}}.$$

An analogous condition holds for the second good. The first-order conditions are characterized by the consideration of the quality-correction terms $Q_i := (1 - \lambda)\Sigma_h \partial r^h / \partial q_i$. Once again, λ is a scaling parameter which is zero for marginal-cost prices and unity if we have an unconstrained monopoly. Therefore, the quality-correction terms Q_i vanish for the perfect monopolist: the profit maximizer neglects consumer welfare gains, the welfare maximizer takes them into account. As these gains are measured by the negative Q_i 's, we may conclude that the welfare-maximizing regulator behaves like a monopolist who underestimates the marginal quality costs by the sum of the individual rates of substitution between quality and income. This implies a tendency towards higher qualities.

Finally, let us briefly sketch the changes in the optimal qualities if an informationally-demanding regulatory process is applied. The cost function will now depend on the quantities produced, on the quality indicators, on the effort variable and on the cost characteristic. An analogous extension holds for the demand functions. Differentiating the managerial incentive-compatibility constraint with respect to the quality indicators gives *quality-induced incentive-correction terms* (and there are still the usual quantity-induced incentive-correction terms which we have treated in the section on informationally-demanding regulatory rules). The rules on price regulation remain unchanged, although the absolute levels of prices will, of course, change. However, we get new rules on quality regulation. Both in the case of asymmetric information on costs and in the case of asymmetric information on demand, a quality-induced incentive-correction term is added to the marginal quality costs – the regulator considers the sum of the quality costs and the costs that are induced by the information rent of the manager. This holds for both cases of asymmetric information treated in this paper (on costs and on demand).

Conclusion

The most important contribution of the new theories of price regulation has been the accentuation of the information and incentive structures. Imperfectly informed regulators may set wrong incentives for the managers of the regulated public utilities. The new theories of regulation show how to achieve the best possible results if the regulator lacks information. This paper has first presented various simple regulatory mechanisms. Vogelsang and Finsinger's iterative mechanism excels by its minimal information requirements: the regulator only needs information on past quantities, past prices and past realizations of total costs. Shleifer's yardstick regulation links the regulatory rules for some firm to performance indicators of other firms in similar position: since truth-telling does not influence a firm's own regulation, it will not cheat, and the regulator gets all the information he needs. Unfortunately, however, in practical applications of these simple mechanisms the regulated utility will be able to dodge the regulator's intentions by strategic behavior. This has led us to a treatment of informationally-demanding regulatory mechanisms which are incentive-compatible and, therefore, strategy-proof. Unfortunately, however, the regulator must be extremely well-informed if he wants to apply this sort of regulation: except for the actual realizations of a cost or demand characteristic and the effort of the manager, he must be perfectly informed about the situation of the regulated public utility. This must be the main reason why in practice the simple $RPI - X$ regulation prevails. Obviously, it represents an acceptable compromise between not too high information problems and not too high incentives for managerial strategic behavior. Finally, in this paper we have accentuated the importance of quality regulation, which is a rather neglected field in the theory of regulation and also very often in the practice of regulation.

References

- Armstrong, Mark, Rees, Ray and John Vickers (1991), 'Optimal Regulatory Lag under Price Cap Regulation', mimeo: Nuffield College, Oxford.
- Baron, David P. and Roger B. Myerson (1982), 'Regulating a Monopolist with Unknown Costs', *Econometrica*, 50 (4), 911-30.
- Baumol, William J., Panzar, John C. and Robert D. Willig (1982), *Contestable Markets and the Theory of Industry Structure*, rev. edn. 1988. New York: Harcourt Brace Jovanovich.
- Bawa, Vijay S. and David S. Sibley (1980), 'Dynamic Behavior of a Firm Subject to Stochastic Regulatory Review', *International Economic Review* 21 (3) 627-42.
- Boiteux, Marcel (1956, 1971), 'Sur la gestion des monopoles publics astreints à l'équilibre budgétaire', *Econometrica*, 24 (1), 22-40. (English edition: 'On the Management of Public Monopolies Subject to Budgetary Constraints', *Journal of Economic Theory*, 3 (3), 219-40.)
- Bös, Dieter (1991), *Privatization: A Theoretical Treatment*. Oxford: Oxford University Press.
- Bös, Dieter (1994), *Pricing and Price Regulation / An Economic Theory for Public Enterprises and Public Utilities* (Advanced Textbooks in Economics, Vol. 34). Amsterdam: Elsevier, North-Holland.
- Bös, Dieter and Peters, Wolfgang (1991), 'A Principal-Agent Approach on Manager Effort and Control in Privatized and Public Firms', in Attiat Ott and Keith Hartley (eds), *Privatization and Economic Efficiency*. Aldershot, UK and Brookfield, US: Edward Elgar, pp. 26-52.
- Laffont, Jean-Jacques and Jean Tirole (1993), *A Theory of Incentives in Procurement and Regulation*. Cambridge, Mass.: MIT Press.
- Leibenstein, Harvey (1966), 'Allocative Efficiency vs. "X-Efficiency"', *American Economic Review*, 56 (3), 392-415.
- Lewis, Tracy R. and David E.M. Sappington (1988), 'Regulating a Monopolist with Unknown Demand', *American Economic Review*, 78 (5), 986-98.
- Littlechild, Stephen C. (1983), *Regulation of British Telecommunications' Profitability*. London: HMSO.
- Loeb, Martin and Wesley A. Magat (1979), 'A Decentralized Method for Utility Regulation', *Journal of Law and Economics*, 22 (2), 399-404.
- Panzar, John C. (1989), 'Technological Determinants of Firm and Industry Structure', in Richard Schmalensee and Robert D. Willig (eds), *Handbook of Industrial Organization*, Vol. I, Amsterdam: Elsevier, North Holland.
- Ramsey, Frank (1927), 'A Contribution to the Theory of Taxation', *Economic Journal*, 37

- (1), 47-61.
- Sappington, David E.M. (1980), 'Strategic Firm Behavior under a Dynamic Regulatory Adjustment Process', *Bell Journal of Economics*, 11 (1), 360-72.
- Schmalensee, Richard (1989), 'Good Regulatory Regimes', *Rand Journal of Economics*, 20 (3) 417-36.
- Sharkey, William W. (1982), *The Theory of Natural Monopoly*. Cambridge: Cambridge University Press.
- Shleifer, Andrei (1985), 'A Theory of Yardstick Competition', *Rand Journal of Economics*, 16 (3), 319-27.
- Vickers, John and George Yarrow (1988), *Privatization: An Economic Analysis*. Cambridge, Mass.: MIT Press.
- Vogelsang, Ingo and Jörg Finsinger (1979), 'A Regulatory Adjustment Process for Optimal Pricing by Multiproduct Monopoly Firms', *Bell Journal of Economics*, 10 (1), 157-71.