Chapter 4.

Monopoly regulation

Tomás Gómez

Table of contents

4.1. Fundamentals of monopoly regulation................................................................. 3
4.2. Cost of service regulation .................................................................................... 7
  4.2.1. Determining the cost of service................................................................. 7
  4.2.2. The rate base ......................................................................................... 10
  4.2.3. The rate of return ............................................................................... 11
  4.2.4. Strong and weak points ....................................................................... 11
4.3. Incentive-based regulation .................................................................................. 15
  4.3.1. Price cap ............................................................................................... 15
  4.3.2. Revenue cap ......................................................................................... 17
  4.3.3. Mechanisms for sharing earnings and losses ......................................... 19
  4.3.4. Design of an incentive-based regulation scheme ..................................... 20
  4.3.5. Strong and weak points ....................................................................... 22
4.4. Implementation details for price or revenue caps .............................................. 24
  4.4.1. Present value of costs and revenues: the smoothing X ........................ 24
  4.4.2. Regulatory asset base, investment and depreciation ......................... 28
  4.4.3. Calculating the WACC ........................................................................ 32
  4.4.4. Operating costs and benchmarking ..................................................... 34
  4.4.5. Tax, cash flow and profitability ......................................................... 38
  4.4.6. Consumers or companies risks ............................................................ 41
4.5. Quality of service and other issues ................................................................. 43
4.6. Summary ......................................................................................................... 45
4.7. References ...................................................................................................... 47
“For a few decades after September 1882, when Thomas Alba Edison commissioned the first electric power plant on Pearl Street in New York City, the electric power industry was composed of small firms in competition with one another. Then, for many years until barely two decades ago, it has been considered as a natural, regulated monopoly.”

This chapter is an introduction to the fundamentals of natural monopoly regulation, particularly as it applies to utilities providing what are regarded to be public services: electricity, water, telecommunications and gas, although, given the subject matter of this book, the focus is logically on electricity.

Monopoly regulation is justified by the need to prevent a monopolistic provider from overcharging consumers or delivering a service of unacceptable quality, given the loss of economic efficiency this would entail for society as a whole.

When the monopolist is a State-owned company, the State can act as both owner and regulator without clearly separating the two functions. On the other hand, when the monopolist is a private company or there is a separation of functions, the State takes responsibility of establishing the regulation, whose implementation and supervision may be delegated to regulatory commissions. For instance, in the US regulatory commissions were created to control utilities in all 50 states, and were in place throughout the twentieth century. The US also has nation-wide regulatory commissions such as the Federal Energy Regulatory Commission (FERC) and the Federal Communications Commission (FCC), whose mandates cover interstate transactions or services that fall outside the regulatory powers of individual states.

The preceding decade has often been said to have witnessed the collapse of natural monopolies, but this is not entirely true. For while competitive markets have in fact been created for telecommunications, electricity and gas, as well as water supply services, these markets have focused on wholesale production and retail sales of the product in question: information, kilowatt-hour of electric or thermal power, cubic metres of water and so on. However, building and maintenance of the necessary network infrastructures continue to operate as natural monopolies (this is only partly true in telecommunications).

One of the chief characteristics of electrical network infrastructures, which depends on large investments, is that they are “bound” to the physical space where they are located. Indeed, one often cited example that clearly illustrates the inefficiencies inherent in introducing competition in this type of activities is the duplicate expense involved in two competing electricity distribution companies building the same type of infrastructure in the same area to provide the same service. Their networks would be redundant and users would end up paying roughly double the price for the same service.

The first section of this chapter establishes the fundamentals and principles on which monopoly regulation is based and introduces the variables that can be regulated: company revenues, quality of service standards, past and future investment, and the conditions to be met by the exclusive licensee.

The second section gives a detailed description of the procedures normally followed by regulators to establish revenue ceilings and tariffs under the most traditional regulatory
method, known as cost-of-service or rate-of-return regulation. The advantages and drawbacks to this approach are also discussed.

The third section addresses the differences between incentive-based and cost-of-service regulation and justifies the growing acceptance of the former as an alternative to enhance efficiency where electric power transmission and distribution are separate and regulated activities. The two most popular methods, price and revenue caps, are described at length.

Finally, the last section contains detailed information intended to provide an in-depth understanding of the various stages and basic elements involved in a price review process, and other issues that should be regulated, such as quality of service.
4.1. FUNDAMENTALS OF MONOPOLY REGULATION

A monopoly exists when, for whatever reason, a given company becomes the sole supplier of a product or service. Further to the discussion in Chapter 2, markets are the most efficient way of producing and selling goods or services. Under certain circumstances the conditions required for an acceptable level of competition are not present, however. In other words, markets can fail as a result of market concentration, economies of scale, public goods, externalities, incomplete information or transaction costs. Market failures must be corrected by regulatory intervention to ensure the optimal outcome for society. When economies of scale exist, economic reasons can be put forward to explain why a monopoly is the only feasible option to provide a product or service. Natural monopolies may be characterised by one or more of these features: i) economies of scale, ii) capital intensity, iii) non-storability with fluctuating demand, iv) location-specific delivery generating location rents, v) production of essential services for the community, and vi) direct connections to customers. While all these circumstances are present in the electricity industry, some segments of this industry, generation and retail, can migrate from a monopoly to a market structure. The transmission and distribution grids, however, must continue being regulated as a natural monopoly.

An unregulated monopoly would be in a position to charge consumers a price much higher than its production costs, with the concomitant loss of economic efficiency. Consequently, some form of regulation is required to ensure efficiency under these circumstances (Posner, 1999).

The electricity industry has been traditionally dominated by national or regional monopolies subject to price control regulation, with the regulator setting new tariffs from time to time. After the wave of industry deregulation and liberalisation in the 1990s, the generation and sale of electric power to end consumers are viewed as activities that can be conducted under competition, whereas network activities, i.e., electricity transmission and distribution, are considered to be natural monopolies and still in need of regulation. The justification, in terms of economic efficiency, is immediate: quite obviously, allowing two or more companies to build power lines across the same region to supply the same community of consumers with electricity would be both prohibitive and wasteful.

According to economic regulation theory, the monopolistic supplier of products or services regarded to be in the public interest should be prevented from exploiting its market power, either through due regulation or by some other form of control, such as State or public local institution ownership. Yet the same ideal of economic efficiency that underlies the operation of perfect markets –namely, that the profit motive induces innovation, investment in new projects and reduced costs, and competition puts downward pressure on prices–, should also inform the design of efficient monopoly regulation.

Regulators may choose from a number of regulatory variables as tools to reach efficiency objectives (Viscusi et al, 2005). Arguably, the most important of these approaches is the regulation of the revenues from electricity sales that the company is allowed to earn. Such revenues must be sufficient to enable the utility to cover its operating costs and make any necessary investments, while earning an adequate return on the capital invested. In other words, revenues should ensure the company’s medium- and long-term economic and
financial viability, without driving it to bankruptcy. Conversely, such revenues should not be detrimental to consumer interests. Moreover, in the case of an essential service such as electricity, unjustifiably high prices also have an adverse impact on the competitiveness of a country’s industry.

Generally speaking, the (conflicting) objectives that appropriate regulation should pursue in determining the allowable volume of regulated revenues are:

• economic and financial sustainability of the utility.

• productive efficiency, attempting to provide the service or product at the lowest possible cost.

From the standpoint of tariff design for the end users, the following objectives must be borne in mind:

• equity, whereby the receipts from any given community of consumers cover the cost incurred by their consumption, ruling out cross-subsidies among consumer groups.

• pricing efficiency, whereby the amounts charged should be kept as close as possible to the marginal costs of providing the service or product.

• sufficiency, whereby the receipts from tariffs concur with the revenues allowed by the regulator.

Where network companies are privately owned, sight should not be lost of the large, immovable investment required to build the associated infrastructure. If the regulator yields to political or populist pressure to lower prices so far that the return on the company investment is insufficient, the result will very likely be a firm decision on the part of the utility not to invest even the amounts strictly necessary. Regulation should strike a proper balance between operating costs and capital investment. If heavy investments are made at the expense of sufficient staffing, for instance, the outcome is not economically efficient. If, on the contrary, the company skimps on investment to pay shareholders higher dividends, the medium-term quality of supply will be compromised.

Another variable that regulators may draw on to enhance efficiency is the service quality standards that the company is required to meet. In the electricity industry such quality is related to: 1) reliability of supply, i.e., the number and severity of power supply outages; 2) voltage quality, defined as the existence or otherwise of disturbances that may affect the proper operation of apparatus and equipment connected to the mains; and 3) consumer satisfaction with the service standards, for instance time for providing new connections, maintained by the company. All these quality indicators are directly related to operation and maintenance costs, but also to company investment and the quality of the installed infrastructure.

Regulation that excessively encourages cost cutting or lower investment may lead, as noted above, to a gradual deterioration of the quality of the electric power delivered to consumers. For all these reasons, the regulator should explicitly establish performance standards and link them to the revenues the company is allowed to receive.

A variable that sometimes is explicitly controlled by regulators is the investment in new infrastructure proposed by the company in its transmission or distribution grid. This also
has to do with the problem discussed above. The regulator’s primary long-term objective is to ensure that sufficient installed capacity is built to meet the expected demand at suitable levels of quality. This, as noted, is directly related to the rate of return on investment and any deviation, upward or downward, has undesirable consequences. The regulator may attempt to solve this difficult problem by establishing criteria to assess the suitability and necessity of the investments proposed by the company. One example can be the use of network planning models for cost/benefit evaluation of the proposed investments. The increase in regulatory costs entailed in the concomitant control and information gathering must also be taken into consideration, however.

Yet another variable controlled by regulators is the entry into or exit from the business by companies other than the incumbent monopolist. This variable is very important where market regulation is concerned. In the case of electricity distribution utilities, regulation implies an agreement that grants the supplier the right to distribute electric power in a territory exclusively, in exchange for submission to regulatory control. One of the conditions typically imposed on the monopolist is the obligation to supply power to all users, regardless of the associated cost, since electricity is an essential service. Companies may attempt to refrain from servicing high-cost consumers or areas and focus on areas where costs are lower. Regulators must require operators to also provide service of sufficient quality in higher cost remote or rural regions, which must be remunerated accordingly. Alternatively, regulators may establish incentives for other potential suppliers to enter the market: granting territorial franchises to small local cooperatives, for instance.

Practically speaking, the aim is to define the regulatory design that achieves an optimal trade-off between efficiency and quality of service. The following sections review the two most common types of regulation: the traditional method used in the electricity industry for many years, known as cost-of-service or rate-of-return regulation, and an alternative regulatory instrument, which is an extension of the previous one, and has become increasingly popular in many countries, known as incentive-based regulation.

Other less common regulatory methods are also used, and are mentioned here briefly to complete this section. In yardstick competition, the tariffs that a company is allowed to charge are calculated on the basis of the costs declared by other regulated operators in the same business (Shleifer, 1985). Implementation of this scheme requires substantial amounts of comparable information on the characteristics and costs of a number of companies. Each operator’s remuneration is normally established on the grounds of statistical analyses that determine average efficiency patterns for the group as a whole. This gives companies an incentive to lower their costs, since the method for setting revenues is unrelated to their own individual balance sheet. As a result the yardstick value will decline, lowering prices to end consumers. The implementation of yardstick competition also encounters practical problems, however. The regulator must have data for a sufficient number of similar companies, but the companies involved may not all be similar enough to assume that the differences among them are due exclusively to different degrees of efficiency; collusive behaviour may be encouraged, and dissociating revenues from costs may drive some companies into financial straits. The yardstick approach may also be viewed as an input to incentive-based regulation, when benchmarking techniques are used to compare relative efficiency among regulated companies in the same industry, as described in item 4.4.4 below.
A third type of regulation is *light-handed* regulation, exercised by the company itself under regulatory supervision. Under this scheme, the company sets the tariffs to be charged to consumers subject to regulatory approval, which may include mandatory changes in tariff structures and rates.
4.2. COST OF SERVICE REGULATION

In cost-of-service, also known as rate-of-return regulation, the tariffs charged by the utility are authorised and set by the regulator. Tariffs are periodically “re-negotiated” by the regulator and the company in what are known as rate case proceedings. The regulatory process essentially entails two successive stages. The discussion below refers to the way this type of traditional regulation has been implemented in the US and applied to vertically integrated utilities, which may vary slightly from the arrangements in place in other countries.

The revenues the company is allowed to receive (rate level) are determined in the first stage, a process that involves (1) identifying company total costs and investments and (2) establishing the allowed rate of return, to provide the utility with suitable remuneration for the capital invested. Rate cases are based on the data furnished by the company for the preceding accounting period and a forecast of the needs of expenditure for the next control period, and the tariffs established remain in effect for the following period, until the next rate case revision. In the US the rate hearings happen at irregular intervals.

The second stage consists of determining the tariff structure, in other words, of defining the tariff components to be charged to each type of consumer for each cost item. These tariffs are designed to enable the company to collect the revenues allowed by the regulator as calculated in the first stage. For further details on tariff design and calculation, see Chapter 8 of this book.

The rate case or revision process usually comprises the following steps (Rothwell and Gómez, 2003).

• The regulator decides to initiate the tariff revision process because the established period has lapsed, or more frequently, because the utility makes the request of a rate revision.

• After the company submits detailed accounting information and “negotiates” with the regulator, the latter determines the rate of return to be applied to the capital invested and the suitable level of expenses to be covered.

• Finally, tariffs for end users are adjusted to the allowed revenues calculated in the preceding stage. This requires taking into account the energy demanded in the period in question; as this may change with prices depending on demand elasticity, information is likewise required on the latter parameter.

4.2.1. Determining the cost of service

Box 4-1 presents a typical general breakdown of the allowed revenues for a vertically integrated electric utility. The revenues are intended to recover the total costs of providing the service, which must include the generation, transmission, distribution and retail activities. The considered costs essentially include operation and maintenance costs, return on capital, depreciation, and taxes.

Box 4-1. Regulated revenues under cost of service

The following accounting formula represents the balance that the regulator must strike when
reviewing a rate case.

\[ AR = TC = O \& M + DP + s \times RB + TAX - ADR \]  \hspace{1cm} (4.1)

where

- **AR** is the allowed revenues,
- **TC** is the total cost of service,
- **O&M** is the allowed operating and maintenance costs,
- **DP** is the depreciation expenses on the company’s gross assets,
- **s** is the allowed rate of return,
- **RB** is the rate base, a measure of the value of the company’s investment, calculated as its net assets, defined to be its gross assets less depreciation,
- **TAX** is the taxes for which the company is liable, and
- **ADR** is the additional revenue.

The precise definition of terms in the previous formula for vertically integrated electric utilities used in cost-of-service regulation in the US follows (see the American standards on systems of accounts for electric utilities (DOE/FERC, 1973) for further details).

- Operating and maintenance costs: these include the cost of fuel, material and replacement parts, energy purchases, supervision, personnel and overhead.

- Depreciation: the straight line method is generally used. Fixed assets in construction progress are not depreciated.

- Tax: all the taxes for which the utility is liable, i.e., on profit, revenue and property, as well as social security and construction tax (except as relating to fixed assets in progress, since such tax is built into the value of the asset).

- Rate base: this includes net fixed assets (plants, transmission and distribution facilities, other tangible and intangible fixed assets and nuclear fuel, less the cumulative depreciation for all these items), plus current assets (fuel and other material and replacement part inventories, advance payments and deferred revenue, research and development expenses and current asset requirements).

- Rate of return: this is the average weighted interest rate on the company’s long-term financial resources (bonds, debt certificates, shares and preferred shares).

- Additional revenue: this consists of the expenses/revenues deriving from the sale of the company's property, revenues from wholesale energy sales and other revenues not directly related to producing electric power.

Figure 4-1 illustrates how the regulator calculates the allowed cost of service. Allowed costs are the sum of the operational costs (operating and maintenance costs) plus rate base depreciation expenses, plus the allowed returns (rate of return times the rate base), plus miscellaneous (taxes minus additional revenues). As the figure shows, the approved capital investment is added to the rate base and the depreciation is subtracted, and the allowed returns are calculated as the rate base times the allowed rate of return.
Two issues take most of the effort in the negotiations between utility, regulator and stakeholders to determine the allowed revenues for the next price control period: the allowed rate of return, \(s\), and the investments to be included in the rate base. As may be deduced from equation (4.1), ultimately the key figure is the product of these two quantities. Note, also, that the above formula takes account of the allowed O&M expenses, which may not necessarily be the same as the actual expenses incurred. This provides an incentive for the company to enhance its efficiency and constitute a financial penalty for inefficient company management. The simple numerical case example in Table 4-1 presents the list of the accounting items used to calculate the rate of return, in keeping with equation (4.1). An adjustment is made in the calculation of the allowed revenues to improve the allowed rate of return.

| Table 4-1. Example of utility accounts and calculation of rate case, in k€ (Rothwell and Gómez, 2003) |
|-------------------------------------------------|-----------------|-----------------|
| | Anticipated rate case | Adjustment | Rate case after adjustment |
| Revenues | 30 000 | 1 600 | 31 600 |
| Expenses | | | |
| Fuel | 24 000 | | 24 000 |
| Operating | 3 000 | | 3 000 |
| Depreciation | 1 000 | | 1 000 |
| Total expenses | 28 000 | | 28 000 |
| Net operating revenues | 2 000 | | 3 600 |
| Rate base (RB) | | | |
| Assets minus depreciation | 42 000 | | 42 000 |
| Working capital | 350 | | 350 |
| Total rate base | 42 350 | | 42 350 |
| Rate of return | 4.72% | | 8.50% |

Volumetric tariffs are obtained by dividing the allowed revenues by the amount of estimated energy consumption ($/MWh). Volumetric tariffs are common in the US, at
least for small consumers\(^1\). When the resulting tariffs (per MWh) remain unchanged until the next revision, the system encourages the company to cut costs. If it incurs lower operating costs than anticipated in the rate case, it obtains a higher rate of return. Conversely, if costs are higher than anticipated the rate of return will decline. The incentive to lower costs rises with the length of the regulatory period between tariff revisions, a circumstance known as regulatory lag that underlies the incentive-based regulation approach discussed in section 4.3. If mid-term tariff revisions are allowed and the regulator constantly adjusts the rate of return, however, there will be no incentive for the company to reduce costs.

### 4.2.2. The rate base

One important aspect of tariff revision during rate case negotiations is the criterion used by the regulator to calculate the rate base (RB) and determine which investments should be included. The rate base may be determined in a variety of ways, as described below.

- **On the grounds of the original investment**, i.e., the sum paid by the company for its facilities, less the cumulative depreciation. This is also known as the book value of the asset.

- **On the grounds of the cost of reproducing the investment in question today**, reproduction costs, i.e., an estimate of the present cost of rebuilding or re-purchasing the facilities acquired by the company.

- **On the grounds of replacement cost or the new replacement value (NRV)**, i.e., the cost involved in replacing existing assets with new upgraded facilities available on today’s market technology and costs, but which serve the same purpose.

- **On the grounds of the market value of assets**, i.e., the value that the company with its assets would command if sold on the market.

In the US electricity industry, regulators have traditionally used book value to assess the rate base and have focused their efforts on adjusting the rate of return; hence the name *rate-of-return regulation*.

For electricity distribution networks, if the book value is available and the technology is stable, there is no good reason to use another method, so no unnecessary economic risk is created for the company. In case new technologies appear, incentives should be devised for promoting innovation in the new investments. The other methods for rate base estimation can be used when a reliable rate base does not exist to start with, as it has been frequently the case in power sector privatization processes around the world.

\(^1\) Volumetric tariffs are nowadays criticized in the context of pricing distribution and transmission services, because the companies increase their profits by increasing sales. This incentive can act as a barrier for the implementation of energy efficiency and demand response programs. In many US states, revenue decoupling from sales has been proposed as a remedy to counteract this problem. See the website of the [National Action Plan for Energy Efficiency](http://www.epa.gov/cleanenergy/energy-programs/suca/resources.html).
4.2.3. The rate of return

The most common method for defining the allowed rate of return is to calculate the weighted cost of the different forms of financing used by the company, such as bonds or shares traded on equity markets. This manner of calculating the rate of return is known as the weighted average cost of capital (WACC). An example is shown in Table 4-2. The most controversial item is usually the rate of return on the company's own capital, i.e., the remuneration of its equity, which the regulator typically establishes in accordance with the rates set for regulated companies with a similar level of risk: electric power distribution compared to gas distribution, water supply or communications, for instance, or the actual rates of return of other firms, again with an estimated similar risk level. This issue is discussed in greater depth in section 4.4.3.

Table 4-2. Example of calculation of allowed rate of return

<table>
<thead>
<tr>
<th>Capitalisation ratio</th>
<th>Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonds</td>
<td>28</td>
</tr>
<tr>
<td>Quoted shares</td>
<td>12</td>
</tr>
<tr>
<td>Equity</td>
<td>60</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

4.2.4. Strong and weak points

A classical weak point of cost-of-service regulation as the method for setting the rate of return is the so-called Averch-Johnson effect (Averch and Johnson, 1962). A company allowed a rate of return higher than the true cost of capital has an incentive to over-invest, giving rise to economic inefficiency. Conversely, if the rate of return is lower than the cost of capital, the utility will invest very little and its operating costs will rise, likewise generating economic inefficiencies. Further to that reasoning, regardless of whether in the real world the company behaves as predicted by the Averch-Johnson model, this type of regulation requires the regulator to accurately calculate the utility's real cost of capital.

One of the problems encountered in the tariff revision process is the existence of information asymmetries between the regulator and the company. This asymmetry is mitigated by the regulator requesting as much information as possible, and the company, which must do the work involved in furnishing it. The more involved the regulators become in investment planning and operating cost management, the greater is their understanding of the problems they must regulate. The trade-off, however, is the higher cost of regulation, as more specialised personnel and more sophisticated analytical tools are required, and more burden is also placed on the company.

Cost-of-service regulation has been criticised for not furnishing suitable incentives for companies to reduce costs when tariffs are revised frequently, for instance every year or two. Year after year, the company can recover all its duly substantiated costs and accepted by the regulator. The approach described in the following section, known as incentive-based regulation, attempts to correct some of these shortcomings.
Depending, then, on how it is implemented in practice, cost-of-service regulation has advantages and drawbacks.

The advantages associated with cost-of-service when judiciously applied are: (i) in principle, it enables the company to cover its costs, thereby providing financial stability; (ii) the cost of capital is a control parameter for the regulator, and must not be set above what is strictly necessary to sustain an efficient activity; and (iii) it provides for a good balance between optimal investment levels and quality of service, if measures are taken to prevent over-investment in the event the rate of return allowed is too high, by lowering such rate (A-J effect).

On the contrary, when this type of regulation is implemented ineptly, the advantages may be offset by the drawbacks. Some of these are: (i) this type of regulation does not encourage efficiency and in some cases has evolved towards intrusive and legalistic regulation; and (ii) it may encourage over-investment justified by technical arguments and provides an incentive for companies to incur higher costs, ultimately leading to higher prices for consumers.

**Box 4-2. Cost of service in the US electric power industry during the nineteen eighties**

In the US and many other countries with a privately owned electricity industry, some general criteria were followed to determine “fair” electricity rates in the context of vertical integrated utilities. Electric power rates were supposed to (Edison, 1975):

- remunerate electric power suppliers and expenses incurred in providing the service,
- equitably distribute costs among all users, as far as practical given the limitations of metering and similar facilities,
- provide a reasonable return on capital and attract new resources of funding to finance any new facilities needed to cope with demand growth,
- reward service quality and system operating efficiency,
- promote revenue stability over time to facilitate planning for the future, and
- be simple enough to be readily applied by utilities and understood by consumers.

As noted above, ratemaking can usually be broken down into two stages: i) obtaining the revenue requirement (total revenue authorized by the regulator), and ii) formulating the rate structure for each type of user. US regulatory commissions traditionally have focused on determining the total revenues a utility was to receive for providing the service, trying to ensure that it obtained a reasonable (neither excessive nor insufficient) return on its capital. Regulatory commissions also usually addressed the question of the allocation of total costs to different types of users by defining the tariffs for each one: residential, industrial or commercial; nonetheless, some of them did not actively exercise their regulatory powers in this connection.

The revenue requirement determined the extent to which revenues covered operating expenses, provided for a return on invested capital and were able to attract new funding. In practice, determining the cost of service was an extremely complex exercise. The most controversial aspects of this process were as follows (see Gordon (1981) for instance):

---

2 A detailed version of the traditional US power sector regulation in the 1980s is presented in Annex A of this book.
use of actual current or estimated future values (fuel or capital costs for instance) to prevent rates from lagging behind real costs (regulatory lag), a problem intensified by the duration of the regulatory process,

• establishment of a fair return on shareholder equity,

• valuation of assets on the basis of their historic cost (usual practice in the US), replacement cost or at a “fair value”,

• inclusion or otherwise of the fixed assets in progress in the rate base, and

• accounting treatment for the deferred taxes generated when accelerated depreciation methods are applied.

In practice, regulatory commissions have dealt with all these issues in different ways. A discussion on some of the most relevant aspects follows. The regulatory implications of some of these issues are analysed in section 4.4.

The rate base was defined to encompass current assets and net fixed assets, with the latter accounting for the major share of the base. Net fixed assets were evaluated to be the sum of the non-depreciated property used for company operations. The treatment for net fixed assets varied widely from one regulatory commission to the next. The essential questions were: what should be included, when it should be included and at what value.

Standard practice in assessing fixed assets was to use the historic or original value, which was normally much lower than the current replacement value; certain regulatory commissions allowed utilities to use the replacement cost (computed from extrapolations based on the costs actually incurred), or a “fair” value, which each commission established at its discretion.

One particularly controversial area was the inclusion or otherwise of fixed assets in progress in the rate base. When inclusion was not permitted, no provision was made for remuneration for the fixed assets in progress, and the utility was authorised to include the respective interest costs of the capital invested in the works in progress item. If all or part of the fixed assets in progress was included in the rate base, the interest on the part in question was not charged. The trend in the years discussed here was for a growing number of regulatory commissions to allow the inclusion of the fixed assets in progress, to avoid sudden hikes in the rate base, undue interest fund growth during construction and excessive haste in commissioning facilities (Gordon, 1981). A number of different procedures were in use to appropriately account for variations in the rate base throughout the year.

A peculiar American accounting practice known as standardised accounting lay at the root of another variation in the way the rate base was computed. This procedure consisted basically of keeping fictitious parallel accounts used exclusively to determine the company’s tax liability each year: instead of the usual straight line depreciation used in the real accounts, these parallel accounts provided for accelerated depreciation to reduce taxes in the earlier years and increase the liability in subsequent years. This deferred tax was accumulated in a fund. Most American electric utilities used standardised accounting: the effects of this practice on the determination of cost of service and the tax treatment are discussed in greater detail by Suelflow (1973).

One consideration of great practical importance for utilities in the reference period was the time that frequently lapsed between when a rate was calculated and when it was actually applied. This arose as the combined effect of two factors: the practice of calculating rates on the basis of the cost and energy sales figures prevailing when an application for a rate change was submitted, and the relatively long time lapping between that date and when the regulatory commission’s resolution was forthcoming. The procedures in place to alleviate this problem are listed below.

• Use of estimated values to cover the time expected to lapse before the new rates would come
into effect. This practice was still uncommon at the time, however; regulatory commissions were much more prone to allow slightly higher rates to implicitly compensate for this shortfall (Gordon, 1981).

- Clauses on automatic adjustment of rates to accommodate changes in fuel costs. Critics of this procedure sustained that it discouraged companies from seeking less expensive alternatives when fuel prices rose.

- Adjustment of the rate base as new facilities became operational.

The bone of contention in the ratemaking negotiations between electric utilities and regulatory commissions was the determination of the long-term rate of return on the company's financial resources (bonds, debt certificates, preferred shares and shareholder equity: normal shares plus reserves). Given that the average interest rate on debt (bonds and debt certificates) and the average rate of dividends on preferred shares were in fact known in advance, the problem was merely a question of determining the rate of return on common equity.

The general legislation in effect at the time provided that the rate of return on common equity should a) be comparable to that of other investments with similar risks, and b) be sufficient to inspire confidence in the soundness of the company's financial position so it could raise new capital when necessary. Much has been written about this important question in ratemaking (see Kahn (1988) for a detailed discussion). In practice, regulatory commissions would set this rate after hearing the experts' opinions and considering issues such as the method adopted to assess the fixed assets, the estimated lag between the period when rates were to be applied and when they were calculated and the inclusion or otherwise of fixed assets in progress. The model used to quantify the cost of a utility's own capital on the grounds of the risk associated with the various types of businesses is known as the capital asset pricing model or CAPM, discussed in detail in section 4.4.3.

---

3 Supreme Court, Hope Natural Gas Case, 1944.
4.3. INCENTIVE-BASED REGULATION

The basic principle behind incentive-based regulation is the establishment of relatively long intervals between price controls or “rate cases”. In each price control period, which generally covers four or five years, a specific revenue path is defined to create an incentive to lower costs and thereby increase profits. After the regulatory period lapses, all cost components are thoroughly reviewed (in a process similar to the rate case proceedings described in the preceding section). The outcome of this review is a new formula to set revenues or prices for the following regulatory period. In other words, this type of regulation weakens the relationship between prices and costs; in a way, it can be regarded to stand midway between cost-of-service regulation and deregulation with market-defined pricing. There are two basic alternative schemes for incentive-based regulation, the revenue cap and the price cap. These schemes may be combined with schemes for profit- or loss-sharing with users to reduce risk.

Box 4-3. A scheme for illustrating the power of incentives

Laffont and Tirole (1993) introduced a simple model to establish the link between traditional and incentive-based regulation:

\[ TR = (1-b) \times (\text{costs})_{\text{ex-ante}} + b \times (\text{costs})_{\text{ex-post}} \] (4.2)

where:
- \( TR \) is total revenues (ex post)
- \( b \) is a coefficient, \( 0 < b < 1 \), established ex ante, to incorporate allowed costs
- \((\text{costs})_{\text{ex-ante}}\) are the estimated allowed costs (ex ante)
- \((\text{costs})_{\text{ex-post}}\) are the costs incurred (ex post)

The incentive for the utility to reduce costs is inversely proportional to the value assigned to parameter \( b \). Cost-of-service regulation with frequent rate cases can be likened to a low incentive type of regulation in which \( b = 1 \), since the company is allowed to recover all the costs incurred. A number of regulatory schemes with strong incentive mechanisms, in which \( b \) is close to 0, have been devised and include, for instance, (1) tariff freezing arrangements, (2) cost-of-service regulation with long intervals between rate cases, or (3) incentive-based regulation with long regulatory periods.

Incentive-based regulation involves the \textit{ex ante} raising of the proportion of the revenues that the company receives and the \textit{ex post} lowering of the revenues calculated on actual costs incurred. In this way, the key point of incentive regulation is passing the efficiency gains achieved by the company to consumers in the next regulatory period. Regulation involving incentive mechanisms scantily adjusted to \textit{ex post} costs may lead to excessively high or low total revenues or may even jeopardise the feasibility of the company’s long-term business plan. Hence the need for a rate case, or price control review, every few years. Regulatory design should strike a balance between lower short-term costs, global economic efficiency and long-term viability.

4.3.1. Price cap

In the \textit{price cap} approach, a formula is used to set the maximum yearly price that the company can charge for each service provided, for a period of several years. These prices
are adjusted annually to account for inflation minus a correction factor associated with expected increases in productivity:

\[
P_{m,t} = \overline{P}_{m,t-1} \times (1 + RPI_t - X) \pm Z
\]

(4.3)

where

- \( \overline{P}_{m,t} \), is the maximum price that the company can charge for service \( m \) in year \( t \).
- \( RPI_t \), is the annual price variation per unit (retail price index, RPI, or inflation rate) in year \( t \).
- \( X \), is the productivity factor per unit.
- \( Z \), is adjustments owing to unforeseen events beyond the control of the utility, such as natural catastrophes, environmental regulation or tax hikes.

This form of regulation has been used in the United Kingdom, where it is known as “RPI - X”, to regulate telecommunications, gas and electricity network utilities, and in the United States to regulate telecommunications companies, under the term “CPI – X”. The version known as “revenue cap” (see below) is presently used in many power systems worldwide.

Cost-of-service regulation with tariffs frozen for a period of time covering several years can be viewed as price cap regulation with no correction for enhanced productivity.

Price patterns under a price cap scheme for a regulatory period are shown in Figure 4-2. When \( X \) is greater than 0 real prices decline. The larger the value of \( X \), the more will consumers benefit. If the company is able to lower its costs by improving efficiency more than required by the regulator, it also gains. In some situations prices may tend upward rather than downward: where the regulator recognises high investments in a given regulatory period, for instance. At the end of the price control period the regulator will establish the price for the following year, according to the updated estimates of costs, and will set a new value of \( X \) for the next period. The benefit for the consumers during the next period results from the combination of the new price for the next year and the new value of \( X \).

The general RPI or CPI is not the optimum indicator for this purpose because it reflects variations in consumer prices but not in specific industry costs. In practice, mixed indicators found as a weighted average of individual industry (or company) cost indices and the general inflation rate are also used.
In practice, the price cap formula may be applied to: (i) the average price for the company as a whole, after duly weighting each of the services provided; (ii) the average price to be charged to each consumer class; or (iii) the fixed price of each of the terms comprising the end user tariff. The regulator may choose either to (1) allow the company flexibility to design its end consumer tariffs or (2) design the end user tariff itself to prevent cost shifting among consumer classes.

The price cap used to regulate electric power distribution companies in the United Kingdom, for instance, known as the “average revenue” or “revenue yield” cap (see next item), sets the maximum revenue per sales unit. Sales are calculated as the mean amount of energy delivered to the consumers in each class or voltage level. Moreover, providing they honour this limit on unit revenues, companies are free to design their end user tariff structure, albeit under regulator supervision.

In Chile, Peru and Argentina, on the contrary, the regulator sets the maximum prices that can be charged for the various cost items, such as investment, operation and maintenance, customer management, and the adjustment coefficients applicable to those prices throughout the regulatory period. Price caps are included in the resulting formulas for calculating the distribution tariffs to be paid by end consumers. In this case, price caps are translated directly to end-user tariff design.

4.3.2. Revenue cap

Under the revenue cap approach, the maximum yearly revenues the company is allowed to earn for a period of several years are calculated with a formula that makes provision for yearly inflation less a correction factor associated with expected improvements in productivity. These revenues may be adjusted annually in accordance with one or several cost or revenue drivers that are beyond the control of the regulated company, such as the number of consumers, total energy supplied or, in network companies, the size of the network. Provision is also usually made for adjustments to compensate for exceptional events beyond the company’s control.

The simplest expression for a revenue cap is given by:

\[ R_t = R_{t-1} \times (1 + RPI_t - X) \pm Z \]

\[ P_{m,t} = P_{m,t-1} \times (1 + RPI_t - X) \pm Z \]  (4.3)

where
• $R_t$, is the allowed remuneration or revenues in year $t$
• $\text{RPI}_t$, is the annual price variation per unit (retail price index, RPI, or inflation rate) in year $t$
• $X$, is the productivity factor per unit
• $Z$, is any adjustment owing to unforeseen events beyond the control of the utility, such as natural catastrophes, environmental regulation or tax hikes.

Another possible way to express the revenue cap formula is shown in the following equation, in which the cost driver is the number of consumers (Comnes et al., 1995):

$$R_t = (R_{t-1} + CGA \times \Delta \text{Cust}_t) \times (1 + \text{RPI}_t - X) \pm Z$$  \hspace{1cm} (4.5)

where
• $CGA$, is the consumer growth adjustment factor (currency unit/consumer)
• $\Delta \text{Cust}_t$, is the variation in the number of consumers in year $t$.

Another revenue cap formula used by regulators is:

$$R_t = R_{t-1} \times (1 + \text{RPI}_t - X) \times (1 + \alpha \times \Delta D_t) \pm Z$$  \hspace{1cm} (4.6)

where
• $\alpha$, is the economies of scale factor, typically lower than 1, which provides an indication of how much the regulated costs and therefore revenues increase in proportion to a cost driver, represented by $D_t$.
• $\Delta D_t$, is the increment in year $t$ per unit of the selected cost driver (the formula could include just one or several cost drivers), such as units of energy supplied, number of customers, network size or any combination of the three.

Revenue caps set at the beginning of the regulatory period and adjusted yearly in terms of RPI-X only are known as “fixed revenue” caps. “Variable revenue” caps are adjusted yearly by both the RPI-X factor and other cost drivers.

In yet another variation on the revenue cap theme, known as “revenue yield” or “average yield” caps, a yearly cap is set on the average revenue per unit of demand supplied. As in the case of power distribution companies, in the UK, demand can be expressed as a combination of the number of customers and the total energy delivered. For that reason, “revenue yield” caps are sometimes classified as price caps. The revenue yield cap can be formulated as (notation as in Equation 4.6):

$$\frac{R_t}{D_t} = \frac{R_{t-1}}{D_{t-1}} \times (1 + \text{RPI}_t - X) \pm Z$$  \hspace{1cm} (4.7)

In the revenue cap approach, the end user tariffs defined by the regulator on the grounds of company proposals must be designed so the total revenues do not exceed the allowed
revenues. Since receipts may, of course, be higher than the cap in certain cases, an adjustment mechanism should be designed to correct for such deviations. This adjustment mechanism could also take into consideration when actual revenues fell short of the allowed revenues by providing the corresponding compensation to the company next year.

Although the incentives to lower costs are similar in revenue and price cap schemes, increases in sales generate very different effects in the two. Whereas both the price and the revenue yield cap encourage higher sales, revenue caps, depending on how remuneration for sales growth is incorporated in the formula, is always more neutral to this effect. It may therefore be deduced that compatibility between remuneration for the utility and its energy savings or demand-side management programmes can be more effectively attained with revenue caps. Moreover, due to the nature of network-related costs, the greater share of the costs of the regulated network company varies little with demand in the short and medium term. Consequently, revenue caps including appropriate cost drivers duly adjusted for economies of scale are the most popular scheme for price control in this type of regulated businesses. In the US, a good number of state regulatory commissions have implemented “revenue decoupling” from sales when regulating utilities in different forms of revenue caps, providing examples of how this partial decoupling can be designed and work in practice (National Action Plan, 2007).

4.3.3. Mechanisms for sharing earnings and losses

A variation on rate-of-return regulation often used in combination with price or revenue caps is what is known in the literature as the sliding scale. The chief property of this method is that it provides a mechanism for the utility and consumers to share the risk of very high earnings or losses.

**Box 4–4. Sliding scale**

Expressed in cost-or-service or rate-of-return regulation terms, this scheme is based essentially on an adjustment of new rate case prices so that the allowed rate of return in year t, \( s_t \), adopts the following form (Viscusi et al, 2005):

\[
    s_t = s + h(s^* - s_t)
\]

where:

- \( h \) is a constant parameter ranging in value from 0 to 1 and established by the regulator,
- \( s_t \) is the rate of return that the company would obtain in year t as a result of the tariffs set in the current rate case and prior to the application of this adjustment, and
- \( s^* \) is the target rate of return.

If \( h=1 \), tariffs are set in each rate case to ensure the company a rate of return of \( s^* \), i.e., rate-of-return regulation. If rate cases are scheduled at short intervals, the company fails to benefit from efficiency or to lose money for inefficiency. On the contrary, if \( h=0 \), the result is constant or frozen tariff regulation, with all the earnings or losses attributed entirely to the company, i.e. fixed-price regulation. A value of \( h=0.5 \), in turn, would indicate that the profit or loss would be shared by the company and its customers.

Sliding scale ratemaking may also be viewed as a mechanism for adjusting tariffs to maintain the rate of return within a pre-established margin (Comnes et al., 1995). Under
this approach the frequency of rate cases would be reduced by lengthening the regulatory interval. If earnings are so high that the rate of return exceeds the highest value on the scale, tariffs are revised downward. And conversely, if earnings dip below the lowest value, tariffs are revised upward. If the rate of return remains inside the established range, tariffs are not changed and the earnings or losses are attributed to the company. Practically speaking, then, this mechanism divides earnings and losses between the company and its customers.

Such sliding scale mechanisms are sometimes used by regulators in combination with price or revenue caps to protect both company and consumers from the risk of extreme earnings or losses that fall outside what is regarded to be a reasonable margin.

Sliding scale mechanisms may be progressive or regressive. In progressive mechanisms, the part of the profit retained by the company rises with cost savings. For example, a company might receive 20% of the first 5% saved, 40% of the second 5% and so forth. Since cost savings are more difficult to achieve as total cost declines, progressive mechanisms that provide for retaining a higher proportion of the savings constitute more effective incentives for the company than regressive mechanisms.

Examples of sliding scale mechanisms applied to performance-based ratemaking (see the end of next section 4.3.4) for electricity companies in the US can be found in Comnes et al. (1995).

4.3.4. Design of an incentive-based regulation scheme

Designing an incentive-based regulation scheme such as a price or a revenue cap calls for making certain key decisions, addressed in the discussion below (see Navarro (1996) and RAP (2000) for further details).

Definition of the regulatory period

Four or five years is normally a good compromise, as it leaves sufficient time to create incentives for the company to lower its costs (productive efficiency) without running the risk of prices or revenues deviating too far from costs (seeking both financial viability for the utility and cost-of-service efficiency for consumers). If particularly significant changes take place that entail substantial deviations from the initial estimates for the regulatory period, a rate case may be initiated to review the price or revenue formula before the regulatory interval expires. Today there is a discussion on the need of enlargement of regulatory periods for better ensuring the recovery of riskier investments in new clean energy technologies and innovation in energy infrastructures.

Determination of baseline and adjustment parameters

Throughout the review and before the regulatory period begins, regulators must proceed to analyse all the information furnished by the company on past costs and future investment plans. Thereafter they must reach a decision on the costs that the company will be allowed to recover in the following period. This may involve the use of detailed analyses of each cost item and, wherever possible, engineering or econometric models that can also process data on the other companies in the group for benchmarking purposes.
In practice regulators classify costs in several categories.

- **Operating expenses** (OPEX) cover personnel, maintenance and operation.

- Capital expenditure (CAPEX) is related to investments. From the standpoint of the annual allowed cost this heading includes: i) yearly depreciation, and ii) the return on the rate base (RB) or the regulated asset base (RAB), which includes both existing installations not totally depreciated and projected investment during the price control period.

- Uncontrollable costs refer to taxes, upstream fees and other exceptional cost items.

Regulators may use one of two approaches to set the baseline for revenue or price caps and adjustment parameters. These parameters determine how revenues or prices evolve during the regulatory period. The chief factors involved are inflation, productivity and variation in market size: the variation in the number of consumers or energy delivered, for instance. The two approaches are known as the building blocks and the total expenditure (TOTEX) approach. For a fuller discussion see Ajodhia (2005) and Petrov and Nunes (2009).

Under the building blocks approach, the regulator assesses the OPEX, CAPEX, and uncontrollable cost caps separately for each year of the regulatory period. OPEX and investment efficiency can be determined by benchmarking. At the end of the period the regulator monitors both OPEX and CAPEX efficiency gains.

The building blocks approach is used to calculate total maximum allowed revenues directly, which is the sum of each year’s OPEX, CAPEX and uncontrollable costs. A revenue cap formula is calculated as follows: a starting value is determined and adjusted yearly to account for inflation and productivity (X factor) so as to equate allowed revenues to the total costs calculated for the year.

Under the TOTEX approach, separate caps are not explicitly defined for regulated OPEX and CAPEX. Rather, a single TOTEX cap is calculated by the regulator. In TOTEX, a company's productivity factor X for efficiency improvements is usually determined by the regulator by benchmarking. In theory, this approach affords the regulated company greater freedom to make the optimal choice in the trade-offs between OPEX and CAPEX to reduce costs.

Since in practice regulators use a mix of procedures or strategies to define the revenue cap formula, the resulting method can seldom be classified entirely under one approach or the other. Regulators may, for instance, use the building blocks approach to calculate the revenue cap formula and then implement the cap under a TOTEX scheme, whereby the company is assessed on the grounds of total savings.

**Definition of secondary objectives associated with the incentive regulation scheme**

In addition to the primary objective pursued with this type of regulation, i.e., lower costs, other specific aims relating to the characteristics expected of the service provided by the utility may also be sought. These objectives are related to company performance. This type of mechanism is also known as performance-based ratemaking or regulation (PBR), where the company’s remuneration depends on its actual performance in meeting the
specific targets defined by the regulator. Examples of such objectives include improvements in quality of service and consumer satisfaction, universal service for all consumers within the franchise area, reduction of the environmental impact caused by the company’s activities and implementation of programmes in the public interest such as research and development or energy efficiency schemes. The regulatory scheme must include explicit financial mechanisms to reward or penalise the company for reaching or failing to reach the targets set for each objective.

4.3.5. Strong and weak points

As noted above, the main advantages to incentive-based regulation are related to the provision of clear and simple incentives for efficiency through cost reductions. In addition, the amount of information required from the companies and the cost of regulation itself are lower than under traditional cost-of-service regulation with frequent price reviews (Joskow and Schmalensee, 1986). Incentive regulation has also proved to be very useful in countries with scantly developed auditing systems, where State-owned companies were divided and privatised. In such cases, incentive regulation had to strike a balance between company and consumer interests under conditions in which the information available was incomplete.

The implementation of incentive-based regulation schemes also has its weak points, however. The first is the potential for a decline in quality of service. The incentives for companies to lower costs may have an adverse effect on service quality. Facility maintenance and investment costs are directly related to the quality of the service delivered. Consequently, the regulator must necessarily set both quality standards and financial penalties for the failure to meet them.

Excessive concern over the profits that companies may be making may lead regulators to gradually revert to cost-of-service regulation. Such a concern may induce regulators to increase the frequency of rate cases to review costs, ultimately slipping back into traditional regulation and losing the potential benefits of cost reductions.

In keeping with the pursuit of fair profit sharing between customers and the regulated company, key aspects that the regulator should resolve in each price control review are the starting point, the initial price or initial revenue and the choice of both the X factor and the inflation index, which may include a mix of price indicators. All these decisions impact the delicate balance between profit sharing and achieving efficiency and stability in the medium and long term.

Since incentive regulation may lead to laxer cost supervision by the regulator, companies may tend to shift the costs incurred in the non-regulated line of business to their regulated activity, if this is the case (e.g. in an attempt to show that their profits are falling and thereby obtain a larger tariff hike than would be strictly necessary). In addition, monitoring OPEX and CAPEX separately where the requirements for CAPEX are less strict than OPEX introduces the risk of operating costs being capitalised through investment (NAO 2002). One way of tackling this problem is to impose strict rules that require regulated companies to furnish accounting information on standardised forms with a high level of cost disaggregation.
In schemes that allow for company flexibility in establishing the end user tariff structure, mechanisms should be in place to prevent the *shift of costs to consumers with fewer options* or influence over the company: from large industrial to residential consumers, for instance.
4.4. IMPLEMENTATION DETAILS FOR PRICE OR REVENUE CAPS

This section contains a practical description of some of the notions and issues mentioned above that should be considered by regulators during price control reviews when setting caps for the next regulatory period.

Firstly, an alternative viewpoint of the “productivity factor X” will be presented. So far in this chapter, X has been considered as an external input to the process, which is obtained by knowledge about the potential efficiency improvements of the considered company or comparison with similar firms. In practice, this use of X may be possible with some kind of companies or with specific activities within a company (e.g. operation and maintenance costs), but not with others (e.g. investment costs). Therefore, the use of a global productivity factor X encompassing the diversity of costs that a company has to incur to deliver a product might be considered impractical.

An alternative view focuses on the estimation of the efficiently incurred costs that the company will incur during the next control period. Each cost item (e.g. capital costs of existing and new investments, operation and maintenance costs, costs of R&D, uncontrollable costs) is individually proposed by the company and acknowledged by the regulator, who takes into account the potential gains in efficiency that are possible in that particular cost component. This is how “productivity improvements” are accounted for in this approach. No “X factor” is used so far. Once the estimated costs for all items have been obtained for all years of the price control period, the net present value of the costs is computed for the initial year using a convenient discount rate. Then equation (4.4) is employed to determine a trajectory of allowed revenues that has the same net present value as the estimated efficient costs. In this approach the role of X is reduced to be an “adjustment” or “smoothing” factor, whose mission is to make sure that the net present values of the streams of costs and revenues for the duration of the price control period are exactly the same. This is the viewpoint that will be adopted in what follows in this section.

4.4.1. Present value of costs and revenues: the smoothing X

Some basic cost definitions were given in Chapter 2. Now we introduce the concept of present value of costs and revenues, i.e., the change in value of money with time. In rate cases, present value calculations are used to determine the revenues required to cover the utility’s expected costs (including a suitable rate of return on assets). This constitutes an economic, as opposed to a financial, analysis of cash flows in the years comprising the regulatory period.

Under cost-of-service regulation, each year’s revenues should equal the sum of the annual operating costs, the rate of return on net assets during the year and the annual depreciation on gross assets (see Equation 4.1).

Assuming that both costs and revenues for the year materialise at year end, all these amounts should be discounted at the beginning of the period to calculate the present net value.
Net assets, in turn, should be re-calculated each year by adding the new assets resulting from investment made during the year and subtracting the annual depreciation on gross assets in service.

The table below contains a sample calculation of revenues in cost-of-service regulation for a price control of two years duration. Estimated costs and allowed revenues for years 1 and 2 are discounted at the starting point (i.e. at the beginning of year 1).

Table 4-3: Present value of utility costs and revenues (Green and Rodriguez-Pardina, 1999)

<table>
<thead>
<tr>
<th>Value</th>
<th>Item</th>
<th>Start</th>
<th>End-year1</th>
<th>End-year2</th>
<th>Key to columns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net present value of estimated incurred costs (2 year example)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest rate</td>
<td>10%</td>
<td></td>
<td>r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>Operating costs</td>
<td>50</td>
<td>48</td>
<td>OC₁</td>
<td>OC₂</td>
</tr>
<tr>
<td></td>
<td>Depreciation</td>
<td>10</td>
<td>10</td>
<td>D₁</td>
<td>D₂</td>
</tr>
<tr>
<td></td>
<td>Investment</td>
<td>20</td>
<td></td>
<td>I₁</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assets</td>
<td>100</td>
<td>110</td>
<td>A₀</td>
<td>A₁=A₀+I₁-D₁</td>
</tr>
<tr>
<td>Return on assets</td>
<td>10</td>
<td>11</td>
<td>Ret₁=r*A₀</td>
<td>Ret₂=r*A₁</td>
<td></td>
</tr>
<tr>
<td>Discount factor</td>
<td>0.90909</td>
<td>0.82645</td>
<td>d=1/(1+r)</td>
<td>d²</td>
<td></td>
</tr>
<tr>
<td>Discounted</td>
<td>Operation costs</td>
<td>45.455</td>
<td>39.669</td>
<td>OC₁*d</td>
<td>OC₂*d²</td>
</tr>
<tr>
<td></td>
<td>Depreciation</td>
<td>9.091</td>
<td>8.264</td>
<td>D₁*d</td>
<td>D₂*d²</td>
</tr>
<tr>
<td></td>
<td>Return</td>
<td>9.091</td>
<td>9.091</td>
<td>Ret₁*d</td>
<td>Ret₂*d²</td>
</tr>
<tr>
<td></td>
<td>Revenue</td>
<td>63.636</td>
<td>57.025</td>
<td>rev₁=oc₁+dep₁+ret₁</td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>Revenue</td>
<td>70</td>
<td>69</td>
<td>rev₁/d</td>
<td>rev₂/d²</td>
</tr>
</tbody>
</table>

In price or revenue cap regulation over a regulatory period of several years, regulators may distribute the utility’s revenues in different ways in each year, as long as its present value for the entire price control period is maintained. Then the X factor can be calculated to smooth the trajectory of revenues across the regulatory period. As a consequence, revenues are set in a fashion that ensures gradual variation throughout the period. This is achieved by the smoothing X, computed by equating the present value of the revenues calculated according to the revenue cap formula (4.4) during the whole regulatory period to the present value of the allowed costs calculated year by year. The outcome is more gradual variations in prices and the corresponding tariffs.

Table 4-4 illustrates the calculation of the smoothing X.

First the regulatory asset base (RAB) is updated yearly taking into account allowed investments and depreciation. All the figures in red are input data. Next, the rate of return or WACC is set by the regulator. Note that the rate of return is expressed in real values, i.e. no explicit provision for inflation is taken into account in this exercise, and the WACC is a before tax figure, i.e., the revenues that the company would obtain before paying taxes are taken into consideration. This subject is explained in greater detail in
the following items. In this simplified example the demand has been assumed to be constant during the entire price control period.

The regulator then sets the allowed operational costs for each year of the regulatory period. In our example these costs decline over time due to the regulator’s efficiency requirements, i.e. an ad hoc productivity factor for OPEX that is estimated by the regulator\(^4\), and are also expressed in real terms, i.e., net of inflation.

Each year’s total allowed costs are the sum of the return on assets, plus yearly depreciation, plus the allowed OPEX.

In this exercise, the smoothed revenue requirements are calculated in such a way that the average price in each year, revenue \( R_t \), divided by the estimated quantities delivered, \( D_t \), can be expressed as shown below, where the global smoothing factor \( X \) is given in percentage and the initial value in year 0, \( R_0/D_0 \), is known:

\[
\frac{R_t}{D_t} = \frac{R_{t-1}}{D_{t-1}} \left(1 - \frac{X}{100}\right)
\]

This type of cap was referred to as the “revenue yield” cap in item 4.3.2. The inflation rate was not factored into the above equation because this exercise was performed with real costs and prices. Nonetheless, including inflation in all the cost items is straightforward.

The \( X \) factor is calculated with an Excel spreadsheet using the “goal seek” tool, by equating the present value of costs to the present value of revenues and taking the WACC set by the regulator as the discount factor. The computed \( X \) can be considered to have a double meaning. On the one hand, it is the adjustment factor that makes possible that the net present values of costs and revenues are equal; on the other hand, \( X \) can be seen as a global productivity factor.

Note that, while OPEX declines by around 4.5 % yearly, the CAPEX rises in keeping with the investment pattern approved by the regulator, causing the RAB to climb progressively. As a result, the total productivity factor, at 0.55 %, is much lower than 4.5 % but still positive, signifying a reduction in average prices in real terms.

\(^4\) Not to be mistaken with the global smoothing \( X \) factor in equation (4.4), which is the outcome of table 4-4.
Chapter 4. Monopoly regulation

One final observation that should not be overlooked is that sales or the quantities delivered by the company remain flat throughout the regulatory period. In other words, projected investment and operating costs are calculated on the assumption of constant sales. In revenue yield schemes, revenues rise with sales. Variable revenue cap formulas in which economies of scale are taken into consideration call for explicitly calculating the rise in investment and operating costs with increasing sales volumes or selected cost drivers. In other cases the regulator makes direct estimations of the diverse costs for each year of the regulatory period without recurring to prescribed formulas. The reader is invited to experiment with this more realistic problem with a spreadsheet similar to the one used in Table 4-4.

A critical issue when designing price or revenue caps is how to connect prices or revenues between two consecutive regulatory periods. The incentives for the company to lower costs are greater when prices move gradually from the final year of the previous regulatory period to the target point in the last year of the new regulatory period (option 1 in Figure 4-3). This approach allows the regulated company to retain the benefits stemming from cost reductions achieved in the previous regulatory period for a longer time. Exceptionally, if the profits expected under this scheme are regarded to be too high, the average prices at the beginning of the new period may be changed in one step (option 2 in Figure 4-3). Observe that under this latter option the regulated company would have...
a disincentive to be efficient at the end of each regulatory period because the achieved efficiency gains would be quickly passed to consumers as a price reduction.

Figure 4-3. Projected costs and prices: smoothed revenue pattern (Green and Rodriguez-Pardina, 1999).

4.4.2. Regulatory asset base, investment and depreciation

The determination of the regulatory asset base (RAB), i.e., the company’s net assets, is of crucial importance in the tariff revision process.

As noted earlier, the RAB is updated year by year by adding investment and subtracting depreciation. A key issue in rate cases or price controls is the determination of the RAB at the beginning of the regulatory period and the criteria for inclusion of new assets in the RAB.

The RAB for existing facilities

The first step in RAB assessment is the decision on how to evaluate existing assets. This is necessary when a comprehensive method, such as the one described here, is implemented for the first time. Also if the adopted procedure to update the existing assets at the end of a regulatory period requires a complete evaluation (see below). This evaluation is typically done during deregulation and restructuring processes entailing, for example, the separation of regulated and competitive activities, the unbundling and privatisation of State-owned companies or the unbundling of vertically integrated enterprises into separate business units.

The four most popular methods to determine the initial RAB or, if needed, subsequent RABs, as mentioned earlier, are: book value, reproduction cost, replacement cost, and market value.

*Book value* is reliable, providing procedures are in place to audit the general accounts and compare the entries to the aggregated and disaggregated investment in facilities. This criterion usually yields stable values that ensure a suitable rate of return while preventing prices from spiking and the company from earning overly large profits. It eliminates a major source of regulatory risk.

*Reproduction cost* is an estimate of the current cost of a replica of the assets considered. For instance, the reproduction costs for distribution assets are the current costs of building the present distribution grid, i.e., the electric lines and substations installed by the company in the past. The use of reproduction costs instead of the original cost rate base constitutes an attempt to assess capital costs at the current rather than the historic values. It creates a significant regulatory risk for network companies, but it might be
appropriate in case of privatization and unbundling of former state-owned companies where no reliable accounting data are available

*Replacement cost* refers to the cost of replacing current assets with new assets featuring the same functionality but updated technology. To use the distribution grid example, the replacement value of the lines and substations in a given region would be calculated by assuming that the lines and substations installed would be replaced by new lines and substations perfectly planned to meet the expected demand using the most efficient technologies currently available. This method is used in a number of Latin American countries, where the NVR (new replacement value) of the distribution network is calculated at the beginning of each regulatory period with the *model firm approach* (Rudnick and Raineri, 1997).

The use of the replacement instead of the reproduction cost is in line with the economic theory that propounds using long-term marginal or incremental costs when pricing a service or product, from a forward-looking rather than a backward-looking perspective (Kahn, 1988). While in theory this is the most economically efficient way of appraising the value of facilities, it may lead to greater price fluctuations from one period to the next, for it introduces technological innovations and price oscillations in the source materials more abruptly. It should be noted that in case of network infrastructures, electricity or gas, due to the long life of those installations that will be on the ground for ever, this method can create an unacceptable regulatory risk for investors.

Finally, in privatization processes, if the RAB is to be assessed from the *market value* of a privatised company, that market value must be equal to the value of the RAB to be calculated, multiplied by the rate of return established by the regulator for this type of utilities. This so-called “market value” is therefore the direct outcome of a regulatory decision. This may lead to a figure that differs substantially from the book value, giving rise to what is known as stranded assets or the opposite, a valuation above the book value. In any privatisation process, before selling companies, the regulator, subject to State-owner agreement, can establish sufficiently attractive rate schedules for the years ahead to command a selling price that satisfies State coffers. A balance must always be struck, however, between State and consumer interest, ensuring that the latter will not be overburdened with high tariffs in exchange for large immediate gains for the public treasury.

**Inclusion of new assets in the RAB**

Additional assets are routinely required for three main reasons: 1) to deliver electricity to new customers or to meet growing demand, 2) to replace aged, deteriorated or obsolete facilities, and 3) to improve service quality or comply with new legal or environmental requirements.

As explained earlier, price or revenue caps are applied *ex ante* to the years in the regulatory period to encourage efficiency in controllable OPEX and projected investment. If the company betters the OPEX target during the regulatory period, the extra cost reductions should clearly enhance its earnings. The allocation of the benefits of efficient investment, however, is less obvious and controversial. Moreover, the subject is closely related to the mechanics of including new assets in the RAB.
When establishing a pattern for expected investment *ex ante*, the assets involved are included in the RAB, and the value of the allowed CAPEX is adjusted accordingly. This is the value that ultimately determines the allowed revenues for the next period (see exercise in Table 4-4).

The *ex ante* inclusion of investment in the RAB generates information asymmetries between the regulator and the company. Utilities normally have no lack of estimates, forecasts and projections on the amounts and characteristics of the infrastructure that will be needed to meet the needs created by expected market growth under different scenarios. Companies have an obvious incentive to over-estimate requirements to raise their RAB and with it their allowed revenues. Company proposals in this regard must be critically re-assessed by the regulator, drawing from the necessary expertise, which should include investment cost benchmarking. The UK regulator has established a sliding scale mechanism to incentivise power distribution companies to make accurate projections for future investment. The aim is to avoid gaming and allow companies to retain part of the benefits gained from efficient investment (OFGEM, 2004), (Joskow, 2006). This mechanism is discussed in detail in chapter 5, on the regulation of the distribution business.

Two situations may arise during the regulatory period: 1) the company invests less than initially estimated in the RAB on which its allowed revenues are based, consequently earning higher profits, or 2) it invests more than planned, in which case it makes a loss.

In pure incentive-based regulation, in the absence of specific circumstances beyond the company’s control that would justify deviations from the initial plans, such as unexpected demand growth, changes in legal requirements or similar, a pure *ex ante* approach is adopted. The regulated company keeps all or part of the earnings or losses (see for instance the profit and loss-sharing mechanisms implemented in the UK discussed earlier (Joskow, 2006)). Further to section 4.3, in incentive-based regulation the strength of the incentive to reduce costs and enhance efficiency declines with the frequency of *ex post* cost reviews. Indeed, if such reviews are conducted yearly to include company investments in the RAB, the *de facto* result is cost-of-service regulation in which the incentive for efficient investment is very low.

The challenge is to be able to distinguish between intentionally postponed investment and genuine cost reductions. To prevent potential gaming in the form of postponing planned investment, regulators who are not in a position to monitor each new facility individually, such as in distribution grid companies, must monitor output variables, including quality of service indicators. The company’s revenues should be lowered if the quality of supply does not improve as expected or even declines due to underinvestment.

When significant individualised investment, such as in transmission assets, for instance, is not undertaken within the expected time frame, some regulators have implemented an *ex post* review known as the *trigger approach*. The company is allowed extra revenues when the investment becomes operational or is penalised when it is delayed with respect to when the asset in question was expected to come on stream (Alexander and Harris, 2005).
As noted earlier, in practice most regulators adopt an *ex ante/*ex post* approach. See Alexander and Harris (2005) for other alternatives. Under this approach the assets approved in the *ex post* review are the ones that replace the assets included in the RAB *ex ante*. The same two situations as described above are possible here: 1) the actual investment made by the company is smaller than initially included in the RAB, in which case the *ex post* review is straightforward, resulting in the inclusion of the actual investment in the RAB; or 2) the company invests more than initially forecast in the RAB, in which case the *ex post* review must include a detailed investigation into whether the investment decisions were adopted prudently. If that investigation deems some of the investments to be unnecessary, the respective assets may either not be included in the RAB or deemed to be eligible for depreciation expenses but not rate of return. The investments found to be necessary and efficient, by contrast, are included in the RAB for full recovery (Alexander and Harris, 2005).

Another issue that must be addressed is *ex post* review timing. Three options can be considered.

- At the end of the regulatory period: assuming regulatory periods of five years, the incentive for efficient investment would be 5 years for facilities coming on stream in the first year of the regulatory period and 1 year for facilities commissioned in the final year.

- At the end of the next regulatory period: the incentive would range from 10 to 6 years.

- On a rolling basis: a 5-year incentive would be established for any investment irrespective of the year when it is actually made.

The longer the time lapsing between the inclusion of an asset in the RAB and the *ex post* review, the greater is companies’ incentive to lower actual investment costs. Consequently, the second of the three options listed generates the most powerful incentive, followed by option 3. See Alexander and Harris (2005) for numerical examples and a fuller discussion of other alternatives.

**Depreciation**

Assets may be depreciated by a number of methods (Green and Rodriguez-Pardina, 1999), two of which are defined below.

- The annuity method: a flat annual charge is applied throughout the life of the facility to recover the capital invested plus the return on investment. The constant annuity consists of the repayment of the principal (amortisation or depreciation itself) on the one hand and the return on investment (interest or rate of return on capital) on the other. The amount of amortisation each year is computed so that the sum of amortisation plus return on the remaining capital is constant.

- The straight line method: the yearly depreciation expense remains constant throughout the life of the facility, and therefore the total annual charge, including the return on investment, declines over time.

The two methods are illustrated in Table **4-5**.

\[ \text{Table 4-5:} \]

\[ \text{Depreciation Methods} \]

\[ \begin{array}{|c|c|}
\hline
\text{Method} & \text{Description} \\
\hline
\text{Annuity} & \text{Flat annual charge applied throughout the life of the facility} \\
\hline
\text{Straight Line} & \text{Yearly depreciation expense remains constant} \\
\hline
\end{array} \]

31
If correctly applied, the two methods yield the same present value of revenues to be received by the regulated company (see Table 4-5), although the impact on present and future customers differs. Straight line depreciation would require present customers to pay more than future customers, while under constant annuity arrangements the two groups of customers would be equally impacted.

Accelerated depreciation practices are also common. In accelerated depreciation either the expense is higher in the early years or the depreciation period is shortened. Like the other methods, accelerated depreciation would not affect the present value of consumer repayments, although today's customers would pay more than tomorrow's (see Table 4-6).

Moreover, depreciation timing also affects the amount of corporation tax paid by the company, for as a deductible expense, it lowers the net operating income on which the tax liability is computed (see item 4.4.5). With accelerated depreciation the company could pay less tax in the early years but more in the final years of the depreciation period. In such cases the regulator should determine how to pass tax savings on to customers (Green and Rodriguez-Pardina, 1999).

In principle, it is up to the regulator to establish both the service life of facilities for the intents and purposes of depreciation and the depreciation method (the straight line method is the most common).

### 4.4.3. Calculating the WACC

Since a company can use both debt and equity to finance its investments, its cost of capital is a weighted average of the interest rate on debt and the expected rate of return on equity. This average rate is known as the weighted average cost of capital WACC.

\[
WACC = \left[\frac{\text{Debt}}{\text{Debt} + \text{Equity}}\right] \times R_{\text{debt}} + \left[\frac{\text{Equity}}{\text{Debt} + \text{Equity}}\right] \times R_{\text{equity}} \quad (4.10)
\]

For example, if 40% of a company's capital is debt and 60% equity, and the interest paid on debt \(R_{\text{debt}}\) is 5% and the cost of equity is 8%, the WACC is 6.8%.
As discussed above, the total rate of return set by the regulator during price or revenue
revision processes is a crucial variable and has a direct bearing on revenues and expenses.
The rate of return (WACC) determines both the average remuneration for the company’s
capital and it can also be used as the discount rate, \( d = 1 / (1 + \text{WACC}) \), applied to find the
present cost of projections when calculating allowed revenues throughout the regulatory
period (see Table 4-4 in item 5.4.1).

The interest rate on debt is usually lower than the rate of return on equity, since
shareholders are more exposed to the financial failure of the company than the lenders.
Then the above formula would appear to infer that raising the percentage of debt in the
company’s overall capitalisation would lower the WACC. Increasing debt above certain
limits, however, also raises the likelihood of company bankruptcy and therefore the
interest rate on debt, since the risk of default of the firm increases intolerably. Since risk
rises with the debt/equity ratio, lenders will demand a higher rate of return.

The cost of equity can be estimated from securities market information on similar
companies or industries. The company’s cost of capital is thereby quantified as a whole,
including different types of business and different levels of risk. The most popular model
used by regulators that quantifies the cost of capital on the basis of the risk associated
with different lines of business is known as the capital asset pricing model, CAPM.

CAPM is based on the assumption that the rate of return for any activity is equal to:

- the rate of return on risk-free assets in the economy in question, regarded to be the
  amount received by investors placing their money in the safest financial assets, typically
  State bonds (this was correct before the financial crisis), computed as the average for the
  last few years for long-term rates, to establish a basis consistent with the life of the
  company’s assets,

- plus a risk premium based on the degree to which the asset tracks the securities market:
  in other words, the specific additional risk associated with the asset over and above the
  average market risk (different methods are used to appraise debt and equity).

The risk premium applied to company equity is assumed to be proportional to a
coefficient \( \beta \), which represents the volatility of the value of the company’s financial assets
(shares) compared to average market volatility (see Rothwell and Gómez (2003) for
further details).

\[
R_{\text{equity}} = R_f + \beta \cdot (R_m - R_f)
\]

where:

- \( R_f \) is the risk-free rate of interest
- \( R_m \) is the expected return on an efficient market portfolio.

Since these rates may be expressed in nominal or real (net of inflation) terms, the first
step is to decide whether nominal or real WACC is to be used. The following expression
indicates the relationship between the two.

\[
(1 + WACC_{\text{nominal}}) = (1 + WACC_{\text{real}}) \cdot (1 + \text{inflation rate})
\]
In countries with no international securities exchange or which lack sufficient liquidity for the type of industries regulated, or where the regulatory or financial risk is perceived to be high, the cost of equity is adjusted upward to include a country risk premium.

In conclusion, the utility’s average cost of capital is calculated as follows:

\[
WACC = \left[\frac{\text{Equity}}{\text{Debt}+\text{Equity}}\right] \cdot \left( R_f + \beta \cdot (R_m - R_f) + R_c \right) + \left[\frac{\text{Debt}}{\text{Debt}+\text{Equity}}\right] \cdot R_{\text{debt}} \quad (4.13)
\]

where:

- \( R_c \) is country risk
- \( R_{\text{debt}} \) is the cost of debt calculated as the interest rate on corporate bonds, including the country premium.

Finally, as discussed below, in earnings and cash flow calculations the interest paid on debt is corporation tax deductible. Consequently the WACC value must be defined as a before tax or after tax rate. Assuming \( t \) to be the tax rate per unit, the following relationships hold:

\[
WACC_{\text{after tax}} = (1-t) \cdot WACC_{\text{before tax}} \quad (4.14)
\]

\[
WACC_{\text{after tax}} = \left[\frac{\text{Debt}}{\text{Debt}+\text{Equity}}\right] \cdot R_{\text{debt}} \cdot (1-t) + \left[\frac{\text{Equity}}{\text{Debt}+\text{Equity}}\right] \cdot R_{\text{equity after tax}} \quad (4.15)
\]

Note that, economically speaking, this means that the company pays interest at only \( 1-t \) of the before tax interest rate.

**4.4.4. Operating costs and benchmarking**

An estimate of the company’s operating costs is needed to be able to compute the tariffs and establish the price or revenue formula for the next regulatory period.

The point of departure in this exercise is generally the audited accounts of costs incurred in the preceding period and a business plan furnished by the company with projections for all the years in the next regulatory period. The company must also break this information down as far as possible to show the different cost items by: activities (facility maintenance, delivery of supply to new users, repair of equipment and facilities and new infrastructure); categories (labour, materials, office material and expendables, energy consumption); type of consumer by service area (residential, commercial, industrial, street lighting).

The problem facing the regulator is how to define a feasible operating cost objective for the period. This target must be able to serve as an incentive for efficiency and the reduction of present costs, as well as to maintain company medium- and long-term sustainability; i.e., the level of efficient costs the company should strive to attain. Once the regulator somehow determines the operating cost objective for the regulatory period, where growth in demand or in number of customers have been already taken into consideration, the result can be expressed by applying a productivity factor, \( X \), to these costs throughout the regulatory period:

\[
OPEX_t = OPEX_{t-1} \times (1 + RPI_t - X_{OPEX}) \quad (4.16)
\]
where

- **OPEX**, is the allowed operating costs in year \( t \)
- **RPI\(_ t \)**, is the inflation rate per unit in year \( t \)
- **XOPEX**, is the productivity factor for the allowed operating costs

If equation (4.16) has not taken market growth (number of consumers or energy delivered) induced variations in operating costs into consideration, this can be replaced by an additional term in the equation.

Benchmarking techniques to compare relative efficiency among regulated companies in the same sector or to compare actual efficiencies against a reference company are used by regulators to calculate productivity factors for operating costs under the building blocks approach or for total costs under the TOTEX approach.

**Box 4-5. Benchmarking techniques**

A company is more productive or more efficient than others if it requires fewer inputs to attain the same outputs, or if it produces more outputs with the same inputs. When the company is producing at its highest ideal productivity it is said to be at its productivity frontier (Coelli et al, 1998). Reviews on the benchmarking techniques used by regulators to assess network monopoly efficiency can be found in Jamasb and Pollit (2003) and Ajodhia (2005).

Figure 4-4 shows inputs and outputs for three companies, A, B and two situations for company C. Companies B and C are efficient because they have already reached their productivity frontier (the rightmost curve). Company A, however, has room to enhance its efficiency, to “catch-up” with the others: its input is the same as B’s but its output lower, while its output is the same as C’s but its input higher. Companies that have already reached their productivity frontier can raise their long-term productivity by adopting newer technologies or innovative processes. This is what is known as the “frontier shift” (the leftmost curve). While productivity calculations should ideally take both effects into consideration, regulators often focus on benchmarking current best practice. The distance between a company’s current productivity and its projection on the frontier is a measure of its inefficiency. The higher the inefficiency, the higher is productivity factor \( X \).

Regulators have a number of different methods or benchmarking techniques from which to choose. These methods involve applying statistical techniques to compare the efficiency of different companies providing the same service in the same or similar countries: electric power distribution or transmission companies, for instance. Correlation analyses are run to compare individual cost items in the various markets. The results of such analyses can be used to define
an average efficiency pattern or identify the most efficient companies (best practice) as models that others should emulate. Benchmarking requires a substantial amount of duly validated information and also serves as the basis for what has been referred to in preceding sections as yardstick competition. The most popular techniques are known as “frontier methods”: data envelopment analysis (DEA), corrected ordinary least squares (COLS), and stochastic frontier analysis (SFA).

The various benchmarking techniques are shown in the flow chart in Figure 4–5.

One traditional, simple and practical method for comparing efficiency among companies in the same industry is based on uni- and multi-dimensional ratios (a weighted combination of unidimensional ratios). It is used, for instance, to calculate costs per customer or unit of energy delivered or the number of customers served per employee.

The input-output relationship in the production process cannot be reflected in this family of methods. Total methods try to solve this problem. Index methods such as total factor productivity (TFP) generate a ratio based on the weighted sums of outputs and inputs. Frontier methods are an analytical method for finding the optimal weighting with which to combine outputs and inputs.

Data envelopment analysis (DEA) is a non-parametric technique that requires no functional relationship between outputs and inputs. Efficiency is defined as the ratio between the weighted sum of the outputs and the weighted sum of the inputs. Each company’s efficiency factor is calculated by solving a linear optimisation problem where the weight factors and the efficiency factor are calculated for that company. The obtained efficiency factor lies between 0 and 1. See Jamasb and Pollit (2003) for details on the formulation of linear programming problems. Companies with an efficiency factor of 1 determine the productivity frontier. For instance, in Figure 4–6, four companies are compared when producing one output by using two inputs. The reported values are represented on the input1 (X1)/output (Y) and input2 (X2)/output (Y) graph. Note that companies 1 and 3 determine the productivity frontier (blue line), i.e. their efficiency factors are equal to 1. They require less input (X1 & X2) than companies 2 and 4 to produce the same output (Y). In this case, the efficiency factor can be obtained graphically. For instance, the company 2’s efficiency factor is calculated as the distance OA divided by the distance O2. Note that efficiency factors decrease with the distance to the productivity frontier.
Parametric methods require an understanding of production or cost functions. Figure 4-7 depicts production costs vs outputs or cost drivers (Y) for several companies. Each point represents one company. The average cost pattern is obtained by fitting a regression line, $C_{OLS}$, to the points using the ordinary least squares (OLS) method. The corrected ordinary least squares method (COLS) defines the efficiency frontier, $C_{COLS}$, in terms of the performance of the most efficient company(ies). For company B, the efficiency factor is calculated as distance EF divided by distance BF.

$C_{OLS} = a + f_1(Y)$

$C_{COLS} = (a - CA) + f_1(Y)$

Stochastic frontier analysis (SFA) is similar to the COLS technique, except that it takes stochastic measurement errors into consideration when estimating the productivity frontier and efficiency factors. Each firm’s distance to the frontier is explained here as the sum of a symmetrical error term (associated with relative efficiency) and a random error term to account for noise in the observations. Known probability functions for the distribution of those errors must be assumed (Jamasb and Pollit, 2003).

Finally when only a small number of companies is involved or when exogenous factors hinder inter-company comparisons, an ideal company may be constructed and taken as a reference. This is known as the norm or reference model method (Mateo et al., 2010). The ideal company is designed to provide the same service as the regulated companies, but at minimum cost. Due to the simplifications normally involved in modelling, the output from reference models cannot be used to directly determine the efficient cost levels that should be allowed. Adjustments must always be made to accommodate actual operating constraints that cannot be included in the model. Nonetheless, the model constitutes an objective on which the company should attempt
to converge and which may serve to compare the efficiency of a set of companies engaging in the same industry. Alternatively, it may be used to compare present and past costs for a given company based on past market growth. In a forward perspective, these models can also be used to calculate how far costs must rise from the reference year to the target year to meet expected market demand growth. Incremental costs, found by dividing the rise in costs by the increased demand met, can be then used to determine allowed revenues and design network tariffs, which for power distributors differ for each voltage level (Roman et al., 1999), (Turvey, 2006).

4.4.5. Tax, cash flow and profitability

Utilities must pay a series of duties or fees and taxes levied on their earnings that affect their cash flow and business profitability. Regulators must take such expenses into account in the tariff revision process when calculating the rate of return and costs that the company is allowed to recover.

The charges to be paid by a utility are associated with the type of business conducted. Distributors must pay a series of local duties, including: municipal tax, property tax, chamber of commerce dues, business licence fees and other local charges. For instance, in Spain, such duties account for approximately 8% of a power distribution company’s gross receipts.

As tax rates cannot be controlled by the company, they should be taken into account when calculating its allowed revenues. Moreover, companies must pay the corporate tax, which for instance can amount to 30 - 35% of the earnings before-tax. As it has been explained in section 4.4.3 regulators take into account this tax when setting the allowed rate of return (equations 4.14 and 4.15).

Box 4-6. Income statement and cash flows

Both local charges and the tax on corporate profit are shown on the company’s income statement, which may adopt the following form:

Table 4-7. Sample income statement

<table>
<thead>
<tr>
<th>Allowed revenues</th>
<th>(-) Operating expenses</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-) Local and other non corporate taxes</td>
<td></td>
</tr>
<tr>
<td>EBITDA (earnings before interest, corporate taxes and depreciation/amortisation)</td>
<td></td>
</tr>
<tr>
<td>(-) Depreciation</td>
<td></td>
</tr>
<tr>
<td>EBIT (Net operating income or earnings before interest and corporate taxes)</td>
<td></td>
</tr>
<tr>
<td>(-) Financial expenses (interest and other debt-related expenses)</td>
<td></td>
</tr>
<tr>
<td>EBT (Earnings before corporate taxes)</td>
<td></td>
</tr>
<tr>
<td>(-) Corporate tax</td>
<td></td>
</tr>
<tr>
<td>Net earnings</td>
<td></td>
</tr>
</tbody>
</table>

The company’s cash flow is calculated from the EBITDA for a given year by subtracting the
interest and principal debt payments, the corporate tax, and the investments.

<table>
<thead>
<tr>
<th>Allowed revenues</th>
<th>(−) Operating expenses</th>
</tr>
</thead>
<tbody>
<tr>
<td>(−) Local and other non corporate taxes</td>
<td></td>
</tr>
<tr>
<td>EBITDA (earnings before interest, corporate taxes and depreciation/amortisation)</td>
<td></td>
</tr>
<tr>
<td>(−) Financial expenses (interest and other debt-related expenses)</td>
<td></td>
</tr>
<tr>
<td>(−) Debt amortisation (debt payment corresponding to the principal)</td>
<td></td>
</tr>
<tr>
<td>(−) Corporate tax</td>
<td></td>
</tr>
<tr>
<td>Cash flow</td>
<td></td>
</tr>
<tr>
<td>(−) Investment</td>
<td></td>
</tr>
<tr>
<td>Net cash flow</td>
<td></td>
</tr>
</tbody>
</table>

The foregoing leads to a series of conclusions on how duties and taxes affect utilities.

- Like operating expenses, duties (including local taxes and similar) must be paid out of allowed revenues. They differ from operating expenses, however, in that the company has no power to reduce or otherwise control them.

- Taxes have a direct effect on the company’s earnings. Depreciation policy affects its tax liability. Depreciating more in the early stages, for instance, implies smaller earnings in those initial years and therefore lower corporation taxes, and consequently higher cash flows.

- Allowed rate of return calculations must be consistent with allowed company earnings calculations. For example, if the WACC defined by the regulator refers explicitly to before-tax income, to be comparable, the company’s actual rate of return must be calculated as before tax earnings divided by the rate base. On the contrary, if the rate of return established by the regulator refers to the after-tax results, the actual rate of return must be calculated by dividing net earnings (i.e., net of taxes) by the rate base (see Section 4.4.3 for the relationship between before and after tax WACCs).

Finally, an example will illustrate how a regulated company’s actual profitability for a given regulatory period is calculated after its revenues are set *ex ante* by the regulator. Table 4-9 gives the results for the regulated company described in item 4.4.1 throughout the 5-year regulatory period, assuming performance to be exactly as predicted by the regulator when calculating the revenue-yield cap. In this example, as earlier, the values shown are net of inflation and company sales are flat throughout the period.
Table 4-9. Actual rate of return of a regulated company under revenue-yield cap arrangements

Table 4-9 shows that company investment and operational costs are as predicted (see Table 4-4), and its financial structure is divided into 60% equity and 40% debt. Note that the 8% before tax WACC is computed based on that 60/40 structure, the 7% rate of return on equity, the 3.8% interest rate on debt and 35% corporate tax. That before tax WACC is the rate allowed by the regulator to compute regulated revenues for each year (see Table 4-4). Net income is calculated as revenues minus operating expenditure, depreciation, and interest payments. Lastly, the actual rate of return on equity is checked to verify that it concurs with the value allowed by the regulator, i.e. 7%.

In the example, the actual rate of return, found as the present value of the after tax profit divided by the shareholders’ equity, proves to be equal to 7%, the value assumed for calculating the 8% before tax WACC. Note that the present value is calculated by discounting at the unknown actual rate of return, using the Excel “goal seek” tool.

As an exercise, the reader may wish to calculate the increase in the actual rate of return assuming that the company reduces its operating costs to a constant value of 50 from the first year of the regulatory period.
4.4.6. Consumers or companies risks

A key issue in implementing revenue or price caps is to analyse how inflation variations, sales growth or recession would affect regulated revenues and company costs, and who, consumers or companies, would bear the risk of such deviations.

Under traditional cost-of-service regulation, the prices or tariffs set during the rate case process remain in effect until the next rate case is studied one or two years later. In this approach, the company’s revenues obviously rise or fall in proportion to sales. If costs rise faster than sales, because inflation is high, for instance, or sales are lower than expected, the company shoulders all the risk, and applies for a tariff revision. Conversely, when costs grow below expectations and/or sales are growing faster than initially predicted, the company benefits.

Viewed from the perspective of incentive-based regulation, price cap regulation is equivalent to cost-of-service regulation as far as the risks of variations in demand or sales are concerned, but with a heavier economic impact because the regulatory period lasts for several years. Under price cap arrangements, however, the general effect of inflation on costs and revenues can be handled directly with the cap formula. In some cases a weighted average of price indicators, i.e., a mix of general consumer price and specific industry indexes, is adopted. Although in general inflation may not correspond to the actual evolution of the cost of the utility.

In revenue cap regulation, revenues can be made to grow by raising only the value of the market variables that have a direct effect on cost increases. One example would be a revenue cap formula for a distributor that takes account of larger revenues collected as a result of a rise in the number of consumers, but not as a result of higher demand for energy by existing customers. As noted earlier, in item 4.3.2, from a practical standpoint this type of revenue cap for distributors is consistent with the implementation of energy efficiency and savings programmes or the introduction of distributed generation.

Finally, the revenues allowed in revenue cap arrangements in general differ from the company’s actual receipts from tariffs. Differences always arise between the expected sales taken as the grounds for the revision process and the company’s real sales during the period the revenue cap is in effect. Companies may have incentives to manipulate the sales estimates submitted to regulators:

- to persuade them of the need for higher investment in infrastructure (the company raises its sales growth predictions)
- to persuade them not to reduce prices or even to allow a price rise (the company lowers its sales growth predictions).

In revenue cap regulation revenues can be easily adjusted ex post to accommodate revenue deviations resulting from differences between ex ante predictions and the actual value of variables beyond the company’s control, such as demand growth or inflation forecasts. In a nutshell, with revenue caps (and in the absence of ex post adjustments) consumers may bear all the risk associated with demand variations, while with price caps that risk is transferred to the company.
The *ex post* approach to acknowledging investments or otherwise for inclusion in the regulatory asset base discussed in item 4.4.2 raises the risk assumed by the company, which is uncertain whether the investment cost will be acknowledged by the regulator. The *ex ante* approach, by contrast, affords companies greater certainty, transferring risk to consumers, who would pay or benefit for any deviations from actual infrastructure needs.
4.5. Quality of Service and Other Issues

As discussed in the foregoing, under cost-of-service regulation companies may have an incentive to over-invest to achieve, for instance, self-imposed quality of service levels primarily on the grounds of technical considerations or with any other excuse, simply because the capital invested earns a generous rate of return. Under incentive-based regulation the quality of service situation is exactly the opposite, however. Regulatory incentives to lower costs and enhance efficiency can lead to deterioration of the quality of service delivered. Quite obviously, both investment in infrastructure and operating costs for maintenance and repair in the event of failures have a direct impact on quality of supply.

In incentive-based regulation, performance parameters or indicators that the company must meet should be defined, along with penalties that reduce its revenues when it fails to do so. This is the alternative provided by incentive- or performance-based regulation to solve the difficult problem discussed earlier in connection with cost-of-service regulation. The solution consists not of monitoring each and every company investment to determine its technical and economic justification, but rather of monitoring the results delivered by the company in terms of both total cost and quality of supply.

Consequently, the regulator must establish a scheme to measure and monitor the company’s performance indicators, i.e., the quality of service offered. Electricity distribution monitoring, for instance, might involve verifying factors such as service restoration time after an outage, response time in meeting requests for new service connections or support for customers filing claims or complaints.

One light-handed regulatory tool to encourage companies to improve quality of service levels is public disclosure of their results. The advantage is that the regulator need not define quality targets or the economic implications associated with failure to meet them. Regulatory costs are therefore low. This type of regulation may be insufficient, however, if specific problems need to be solved or when public opinion fails to influence utility behaviour.

Other stronger mechanisms for regulating quality of service are based on penalties and incentives.

Penalties are established when the company fails to comply with the individual quality standards set by the regulator. Here, the company must pay a penalty to consumers for poor quality service, which may be measured in terms of the number and severity of supply outages in a year or failure to respond in a pre-established time to a request for a new connection to the grid, for instance. Such penalty schemes are based on individual user quality monitoring systems.

Furthermore, integrated price quality regulation has been implemented in several countries to regulate network infrastructures such as electricity distribution. A utility’s yearly revenues may be increased or reduced by a certain percentage, for example, depending on the degree of compliance or non-compliance with the quality standards established for given areas or for the company as a whole. In practice this entails inclusion of a $Q$-factor or a quality incentive term in the RPI formula. In this case the regulator sets target values for specific system level quality indices, such as customer minutes lost or percentage
of customers with a certain number of outages. The utility’s measuring and monitoring systems must, moreover, be open to regulator audit.
4.6. SUMMARY

This chapter discusses the fundamentals of and most common methods for regulating monopolies, with specific reference to their implementation in connection with network industries, in particular electricity distribution companies.

The essential ideas set out in this chapter are summarised below.

• Traditionally, the most common regulatory method is what is known as cost-of-service or rate-of-return regulation. From time to time, every year or two, the regulator analyses the company’s costs, assets and investments and establishes the new revenue requirement for the following period. Such revenue requirement enables the company to cover its operating costs and asset depreciation expenses as well as to earn a rate of return, set by the regulator, on its rate base or net assets. The regulator usually also determines the tariffs for the different kind of consumers, from whose application the revenue requirement is recovered.

• Some drawbacks can be identified in cost-of-service regulation: 1) it provides no incentive for lowering company costs, since the regulator tends to acknowledge all costs incurred; 2) the rate of return is generally set high enough to constitute an incentive for companies to invest more than is economically optimal; and 3) information asymmetries between the regulator and the company are more difficult to manage adequately.

• To mitigate such problems, incentive-based regulation methods are becoming more and more popular applied to regulate network industries. The chief characteristic of this type of regulation is that the tariffs or revenues the company is authorised to charge or receive are kept in place for longer intervals, typically 4 or 5 years. This provides an incentive for the company to lower its costs and be more efficient than under cost-of-service regulation. The resulting efficiency improvements are considered by the regulator in the following price control and therefore the consumers also benefit. The two most commonly used methods are price caps and revenue caps.

• These methods set prices or revenues that the company may charge or receive throughout the regulatory period with a formula with a yearly adjustment factor known as \((RPI-X)\).

• The chief difference between price and revenue caps is that, under price caps, any increase in sales leads to higher revenues; i.e., costs are assumed to increase proportionally with sales. Under revenue caps, by contrast, receipts do not rise in direct proportion to sales, but only in keeping with the selected cost drivers and in the proportion established. Price and revenue caps are frequently applied to regulate transmission or distribution network companies, which are characterised by high fixed investments and whose costs do not normally depend on the intensity of use of the resulting assets.

• Incentive-based regulation that induces companies to lower costs in operation or investment may lead them to do so at the expense of service quality. For this reason, such regulatory mechanisms go hand-in-hand with what is known as performance-based regulation, which not only regulates revenues but sets objectives to be met by the
company. The most important of these objectives are the quality standards that the company must meet. Regulators may establish *system level quality targets* to verify average company performance or *individual standards* to guarantee a minimum quality to each consumer. In both cases, when the company fails to comply with the established standards it is financially penalised by a reduction in its allowed revenues.

- In each *rate case* or *price control* procedure, regulators must project future efficient costs with which to align allowed revenues. Some of the key elements of this process are:
  
  - calculation of efficient operating costs: assessing company cost-efficiency on the grounds of reported costs with benchmarking techniques and/or norm reference models.
  
  - inclusion of assets in the regulatory asset base (RAB): updating the RAB with yearly variations to accommodate asset depreciation and projected allowed investment.
  
  - establishment of an allowed rate of return evaluated on the basis of the WACC, which takes the cost of both company debt and company equity into consideration.
  
  - acknowledgement of other non-controllable costs such as charges, duties and taxes that should also be included when calculating allowed revenues.
  
  - calculation of the X factor, as the factor that equates the net present value of allowed costs, including return on capital, to the net present value of allowed revenues during the regulatory period.

- Incentive-based regulation, because of its emphasis on efficiency, will fail in promoting innovation activities, which are risky and typically need times for maturity that are longer than the price control periods in incentive-based regulation. Addressing this issue is a current open area of research in monopolies regulation.


Chapter 4. Monopoly regulation

4.7. REFERENCES


