



HAL
open science

Yardstick Competition for Transport Infrastructure Services

Dominique Bouf, Julien Lévêque

► **To cite this version:**

Dominique Bouf, Julien Lévêque. Yardstick Competition for Transport Infrastructure Services. ECMT (Ed.) = CEMT. Round Table 129 = Table ronde 129 - Transport services: Limits of (De-)Regulation = L'offre de transports : les limites de la (dé)réglementation, 2006, Paris, France. pp. 63-108 (English language) - pp. 67-115 (version française). halshs-00092671

HAL Id: halshs-00092671

<https://shs.hal.science/halshs-00092671>

Submitted on 1 Oct 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Preferred Citation:

Copyright, © CEMT. BOUF Dominique, LEVEQUE Julien. Yardstick Competition for Transport Infrastructure Service. In : ECMT. Transport Services: The Limits of (De)Regulation, Round Table 129, Paris. ECMT. 2006. pp. 63-108. (OECD Publications). ISBN 92-821-2345-6.

With the Authors and the publisher of the copyright agreement. The present version in PDF is under the copyright of CEMT / ECMT (<http://www.cemt.org/>).

YARDSTICK COMPETITION FOR TRANSPORT INFRASTRUCTURE SERVICES

DOMINIQUE BOUF

JULIEN LÉVÊQUE

INTRODUCTION

Infrastructure provision, and notably railways infrastructure provision, is greatly evolving, even if the reform movement is not very fast. In a few decades, the railways have changed markedly in the US, Japan, Africa, South America, and have begun to reform in the EU. The common feature of those reforms is to try to transform the former “protected” monopolies in more competitive or at least more efficient companies. In this general movement, the EU took an original way which implies the separation of infrastructure and operation. It follows that infrastructure has to be regulated per se and not as a component of an integrated system.

The hypothesis underlying this report is that yardstick competition might be an efficient way to regulate infrastructure provision, at least in its core activity: maintenance. But to support this hypothesis we need to present yardstick competition quite extensively.

Thus, this report is organised as follows:

- the first chapter presents a brief overview of the regulation problem in order to present yardstick competition in the context of the various ways to regulate firms or utilities;
- the second chapter presents more precisely yardstick competition with a theoretical approach;
- the third chapter provides some examples of use of yardstick competition in various industries;
- the fourth chapter addresses the very point of the report: is yardstick competition possible and desirable for infrastructure provision ?

1. BRIEF OVERVIEW OF THE REGULATION PROBLEM

In this paper, regulation will be considered as the choice of institutions and rules necessary to achieve the policy objectives unattainable by market forces. We present this overview of regulation by addressing first the question of the desirability of regulation, together with the main problems associated to regulation and then we will present the main choices offered to regulators.

1.1. In some cases, regulation is both desirable and problematic

We will first present the main reasons why regulation might be desirable, before going to the main difficulties encountered by the regulators.

1.1.1 *Why regulation could be desirable.*

Given the definition of regulation, it follows that regulation is needed where market inefficiencies can be found. Thus it is understandable to find market power, opportunism, bounded rationality, and externalities among the reasons that make regulation desirable.

1.1.1.1 *Market power*

With increasing returns, the market forces lead to monopoly. A monopoly might be efficient, as there is no reason why it should waste money, but the optimal prices for a monopoly are not coinciding with what would be optimal for the society, which consists, for the common wisdom, in the marginal cost pricing. There have been a lot of discussions and some strong disagreements on the

marginal cost pricing rule, but whatever the good price is, there is no reason why it should coincide with the monopoly price, which only maximize the revenue of the monopolistic firm. Thus it is desirable to regulate the prices of a monopoly or to ensure that this monopoly is contestable (this will be discussed later as options for the regulator).

Regarding the transport infrastructures, we are facing obviously some kind of monopoly. But it is noteworthy to add that parallel competition may occur on certain routes. Notably this is the case for US railways on important routes. This is also the case in the countries where tolled motorways coexist with free highways and thus different routes with different characteristics might be offered. Nevertheless, infrastructure is, by itself, a spatial monopoly. We mean that unless an infrastructure is saturated, the economic rationale is to have only one infrastructure between two points. Hence, unless the access of the infrastructure is free, there is market power.

1.1.1.2 Opportunism

Opportunism can be defined as the tendency for one economic agent to take profit of the occurrence of non anticipated circumstances. More precisely, we can refer to the definition used by Klein and al. (1978): once a specialized asset has been set up, there is a tendency by the user of this asset to appropriate part of the quasi rent created by the existence of this specialized asset, to the detriment of the owner of the asset. If vertical integration is not possible or desirable, some contracts are necessary between the owner and the user(s) of the asset, but those contracts are by nature incomplete and thus there is a need of enforcement, but also arbitration and finally regulation.

Concerning transport infrastructure, we have to distinguish road and rail. Tolled motorways can be operated by the state or by franchised companies. But the prices have clearly to be regulated, because of what have been said of the market power inherent to infrastructure. For railways, in case of vertical separation (between infrastructure and operation), any kind of franchising gives ample room for opportunism. Actually, the contract with the operating company cannot be complete and thus many unforeseen events can provide an opportunity to renegotiate. Likewise, if rail infrastructure is private or a public company with a certain degree of autonomy, any capacity enhancement not foreseen can provide an opportunity to renegotiate access charges.

1.1.1.3 Bounded rationality

As pointed out by Hicks (1935) “The best of all monopoly profit is a quiet life”. This quest for a quiet life and the bounded rationality of the people in charge of various kind of decisions, together with the lack of competitive pressure can lead to a sort of slack resulting mainly in the need of too many inputs (employees, energy, capital inputs,...). Inefficiency is the direct result of a non optimising behaviour associated to the enjoyable situation of monopoly.

For infrastructure, this lack of optimising behaviour can be observed in the main tasks of infrastructure management: maintenance and slot allocation (for railways), with, as a result, a decrease in the capacity offered and an increase of the maintenance cost.

1.1.1.4 Externalities

Externalities are a well known cause of market failures. Transport infrastructures are a source of positive and negative externalities. Regulation is necessary in this case because positive externalities tend to be produced in insufficient quantity and negative externalities tend to be produced in excessive quantities. Maybe one of the main problem is pollution. As public transport is less polluting than private vehicle usage, an inter mode arbitration might be necessary, with some incentives (regulation or taxation). Urban road pricing (UK, Norway,...) provides an example of this kind of incentives scheme destined to reduce the modal share of the most polluting mode. Another example is provided by some (French, Swedish,...) railways infrastructure charging systems, with lower charges for freight trains in order to reduce road transport negative externalities. Low access charges are supposed to

induce a modal shift toward rail transport, considered as less polluting. The effectiveness of the infrastructure charging system in reducing pollution is, of course, dependant of the efficiency of the freight operator.

Network externalities can be found where the whole system is creating more value than the mere addition of its part. It is obvious that infrastructure (road and rail) benefits from network economies. But it implies mostly that the design of infrastructure should be carried out in such a way that good interconnections link the different parts of the network. The main question is, thus, the division of a network in different sub networks and the necessary co-ordination between them. It might concern infrastructure design, maintenance and operation.

1.1.1.5 Public policy objectives

Infrastructure construction and operation can be undertaken with some redistribution objectives. Particularly, it might be considered as a policy objective to provide infrastructure (highways or railways) not far from every city of a certain size. In that case, the design of the infrastructure network is oriented toward regional development and not economic efficiency. This might have some effects on the operation of infrastructures, with for example some cross subsidisation. To that end, regulation is necessary.

Any investment, and investment in infrastructure particularly, is grounded on an inter temporal arbitration. Investing consists in trading present consumption for future consumption. Future consumption is usually compared to present consumption with the help of a discount rate. The governments might have a different discount rate than the private households, because the governments tends to value higher the future, as compared to the private agent. Thus, the State might tend to intervene in the choices regarding infrastructure investment and operation because of its concern about the future. The private companies are more short term oriented, especially if they are listed. The stock value becomes a short term objective, to the detriment of the long term perspective, needed for developing infrastructure. Maybe the tumultuous story of Railtrack in UK is an example of this conflict between long term and short term objectives.

1.1.2 Why it is problematic.

We will address more precisely the options for the regulator in the last part of this chapter. But it is noteworthy to mention that the regulator aims at correcting the market failures previously presented. The objectives might be both income redistribution and efficiency, but we will concentrate on efficiency.

1.1.2.1 Multi product firms

Many firms subject to regulation are producing several products. Furthermore, many firms produce different varieties of goods or services which have not exactly the same characteristics. This concerns both the demand side and the production side. For example a car producer might produce different sorts of car. Very frequently the producer discriminates by charging a higher price on a certain category of product and accepting a lower margin on another variety. Moreover, very frequently some costs are not easily attributable to one variety of product or another.

Sometimes, discrimination occurs as cross subsidisation: some clients of the firms are subsidising some others. This is frequently the case with utilities. Regarding transport infrastructure, if there is an access charge, it is highly possible that with the charging system, some parts of the network are actually subsidising some others.

Thus, regulation of multi product firms might be more problematic because it is difficult to:

- know the cost of each product or activity;
- know what is the degree of cross subsidisation;

- appreciate the extent to which this possible cross subsidisation is consistent with the objectives of the regulator.

1.1.2.2 Capture theory

It is common that interest groups try to influence the regulator, in order to promote their interests. This tendency can evolve in a more damageable form: the capture of the regulatory body (Stigler, 1971, Becker, 1983). By the word capture, the capture theory suggests that the firm(s) or some other stakeholders can strongly influence the regulatory body, even to the point where the regulation would be turned against the theoretical objectives of regulation. For example the regulatory body would protect the firms from competition instead of promoting competition. Another example concerns the cross subsidisation mentioned before. Some groups of customer might capture the regulatory body to benefit from lower prices or better quality.

1.1.2.3 Asymmetric information

Another basic problem of regulation is that the regulator lacks information about the firm regulated. For example in the ancient times, in France, it was not uncommon for some state owned companies to hide precautionary some strategic information. We can wonder whether it is still the case. This asymmetric information concerns both the firms (the costs and the efforts necessary to reduce them) and their environment (the demand, the external factors influencing costs or demand). So, usually two kinds of informational constraints are distinguished:

- the hidden actions, or moral hazard, undertaken inside the firm, for example to reduce costs;
- the hidden information on the firm's identity (its ability to exert a high effort), or adverse selection, known by the firm regulated and not by the regulator.

1.1.2.4 Other regulatory constraints

To design a regulation scheme, Laffont and Tirole (1993) add two more constraints to the informational constraint :

- a) the transactional constraints, which consist in the costly constraints of writing and enforcing the contracts.
- b) administrative and political constraints, which consist in respecting the rules governing administration and politics.

1.2. The options for the regulators

This part of the report owns very much to Laffont and Tirole (1993) and Carlton and Perloff (2000). We will present successively various forms of regulation or deregulation, finishing by yardstick competition. A special emphasis will be given to price regulation.

1.2.1 Government ownership versus privatization

At a very high level of generality, according to Vickers and Yarrow (1988) there is not very much to say about privatization, apart from the fact that "ownership matters". But the effects of government or private ownership depend on the market structure, the type of regulation and incentives, and broadly speaking, all the environment of the firm.

Regarding infrastructure, it seems difficult to privatize rail infrastructure, as the British experiment has shown. That doesn't mean that private integrated companies cannot be efficient. To the contrary, the US or Japan examples illustrate the efficiency of private vertically integrated railways. But it seems dubious that infrastructure alone can constitute a private firm, because infrastructure development might not be consistent with the objectives of a private firm, which encompass, among other things, the maximization of the value of the shares. Nonetheless if infrastructure is state owned,

the infrastructure manager can be private. That might be the case, for example of motorways, which are government owned but may have a private management.

1.2.2 Price regulation

Price regulation consists in directly fixing the price of the monopoly. Numerous theories have addressed this question. We will only present briefly the main approaches with respect to their relevance for infrastructure and their “power”. Laffont and Tirole (1993) define the power of an incentive mechanism by the ability to induce the firm to reduce its costs.

1.2.2.1 Cost plus versus price cap

In a cost plus contract, the firm does not bear any of its costs. The regulator pays to the firm an amount of money corresponding to the average costs, including a “fair” profit. This incentive scheme has very low power. To the contrary, with a price cap contract, the prices are limited by an index usually related to the inflation, and including sometimes some anticipated productivity gains. This kind of contract is extremely powerful, as the firm keeps the benefits of its effort to reduce costs. During the first years of the British railways reform, the access charges of Railtrack were predominantly regulated by price cap. This led to cost reduction, as predicted by the theory, but the lack of incentive to invest in capacity enhancement and safety gave rise to the well known problems. As a matter of fact, any price cap regulation is periodically revised: to take into account the evolution of the costs of the firm regulated. Hence opportunistically, the regulator may benefit from the cost reduction policy of the firm. It follows that the firm may not have strong incentive to reduce costs. This is called the “ratchet effect”.

Apparently, there is a trade-off between an incentive regime which tends to extract the informational rent¹ but is almost totally non cost reduction incentive (cost plus) and a regime which leaves at least part of the informational rent to the regulated firm but is strongly incentive (price cap). In the real world, the regulatory regime is very often something in between the two polar cases.

Regarding infrastructure, the pricing system is very often debated in terms of marginal cost versus average cost. The cost reduction incentive questions are somewhat neglected in favor of welfare maximization questions, considering that the costs are given.

1.2.2.2 Short Run Marginal Cost, Long Run Marginal Cost or Average Cost pricing ?

Since 1844 and the famous paper from Dupuit, the marginal cost has been favored for infrastructure charging. But the question is still debated. Average cost pricing has still some partisans. One of the most appealing argument in favor of average cost, pointed out by Coase (1945) is as follows. Charging to marginal cost gives us no information on the utility to produce the whole infrastructure. In other terms, depending on the shape of the demand curve, we don’t know whether the infrastructure is socially desirable or not, even if the willingness to pay for it covers the marginal cost.

Nevertheless, it is generally agreed that the Short Run Marginal Cost (hereafter SRMC) constitutes the best theoretical solution to the question of infrastructure charging. The principle is rather simple. SRMC enable to run all vehicle for which the additional costs borne by the society are inferior to the utility of this vehicle for the society, supposed to be represented by the willingness to pay. Even if the principle is simple, the question of what kind of costs have to enter in the SRMC is rather tricky. In particular, environmental costs and congestion costs lead to specific problems. But SRMC might provide no incentive to invest, as the congestion is a source of revenue. To remove this

¹ The informational rent is defined more precisely in chapter 2. For now, we can put it simply: this is the rent enjoyable by a firm because the regulator doesn’t know its costs.

disincentive to invest, one might be tempted to charge according to other principle, among them long run marginal costs (hereby LRMC).

LRMC is defined as the cost of an additional vehicle when the infrastructure is optimally adapted to the demand. Another approach, comparable to a certain extent, is simply to charge the long run average incremental cost of expanding capacity where the capacity is scarce. The main difficulty with those approaches is to practically calculate the amount to be charged. Increasing the capacity of an infrastructure segment leads to the question of indivisibilities. Thus, the cost might vary considerably from place to place. To the contrary this charging system leads to charges more stable over time and thus facilitate the establishment of contracts between operators and infrastructure managers (Nash and al., 1999). Long stable contracts may justify specific investments such as rolling stock. So, there are some arguments in favour of long run marginal cost pricing, even if it would deprive the public of valuable services (the services with prices between short run marginal cost and long run marginal cost). But, one of the major drawbacks of marginal costs (short run or long run) stems from the fact that marginal cost pricing doesn't fully cover the costs. Thus, to cover the non allocated part of the full costs, some public funds might be needed. They are not given for free or, in other terms, there is a cost to get some public funds. So, other pricing methods might be used in order to adjust the level of public subsidies at the desired target level.

1.2.2.3 Ramsey principle

Ramsey principle aims at differentiating the charges according to the value attributed to the services. The principle is quite simple. The services able to pay something above the short run marginal cost must not be eliminated. Because they contribute to cover the total cost and they deliver a service socially desirable. The services with high value can pay higher prices. The idea is thus to set up a tariff based on the demand for the services offered. The higher the value placed by an operator (and its consumers), the more the demand will be inelastic. So, in order to have higher tariff for the operator who value the most the services, the tariff must be inversely related to the elasticity of demand for that service. This is called the inverse elasticity rule. Mathematically, the markup over marginal cost is proportional to the inverse of the price elasticity of the demand.

Ramsey pricing provides a useful theoretical guideline. However, it requires a great deal of information. Both marginal cost and elasticity of demand must be quantified with a certain degree of accuracy. The Ramsey rule has also been criticized for its failure to protect captive customers.

There are other pricing schemes, but a full analysis of infrastructure charging would be out of the scope of this report. Nonetheless, one conclusion might already be outlined. Regarding infrastructure, pricing is seen as tool to reach given objectives and not necessarily as an instrument to achieve cost efficiency.

1.2.3 Rate of return regulation

Instead of regulating directly the prices, regulators can use the rate-of-return (hereafter ROR) regulation to control the rate of return to capital used in utilities. This kind of regulation allows the regulated firm to achieve a rate of return on the fixed capital, once the depreciation of the capital stock and the operational expenses have been taken into account.

Thus ROR is defined by:

$$ROR = \frac{pQ - OE - D(K)}{p_K K}$$

where:

Q is the quantity of product (or service) sold at a price p.

OE represents the operating expenses.

$D(K)$ is the depreciation of the productive capital.
 p_K is the average price of capital goods.

The main drawback of ROR regulation is the tendency for the regulated firm to overinvest in capital (Averch and Johnson, 1962). Thus, common wisdom about ROR regulation is that it is inefficient, although it might lead to lower the prices to a level under the monopolistic prices. But regulators in the US are not any more in favor of ROR regulation, although it may be practised in other countries.

1.2.4 Making a market contestable

The theory of contestable market appeared in the US in the late 70s and early 80s (Baumol, Panzar and Willig, 1982). The main idea is that if a monopoly practices monopoly pricing or is inefficient, the threat of a new entrant can compel him to improve efficiency and reduce the prices. Thus, in a market where the incumbent monopoly can be contested, the monopoly rent is reduced to zero. Making a market contestable consists mainly of reducing the sunk costs and reducing the entry barriers. But this theory is grounded on a number of assumptions, which are not very frequently verified, notably the credible threat from entry (which means low sunk costs and no strategic game protecting the incumbent). The contestable market theory is even more difficult to analyze for multiproduct firms and complex price scheme. We will not expand on this important theory because it seems to us inappropriate for infrastructure.

If the infrastructure manager is also the owner of the track, obviously it would not be welfare improving to duplicate infrastructure, even if the new infrastructure management is more efficient.

If the infrastructure manager is not the owner, several characteristics of the activity lead to consider the threat of a new entrant as damaging: the infrastructure manager has to develop some long term contracts with both the clients and the suppliers. The threat of a new entrant might make uneasy the establishment of contracts, to the extent that they can be renegotiated by the new entrant.

1.2.5 Franchising

Theoretical works on franchising or “competition for the market” derives mostly from the work of Demsetz (1968). What we call franchising here is recurrent competitive tendering to allocate the whole market to a single firm for a given period of time. Thus franchising is also called competition for the market, as opposed to competition in the market. A contract is established between the successful bidder and the public body in charge of the tendering process. Regarding transport, franchising has already a long history with the urban transport concessions system in France and the railways operators in the reformed UK rail system.

Franchising certainly appears to provide a way to put under pressure local monopolies, and, this kind of regulation is becoming more and more frequent in Europe (Van de Velde, 2003). A lot have been written on franchising, see for example Preston and Nash (1993), for a theoretical analysis of railways franchising, and White (2000) for an analysis of competitive tendering for London buses. The main difficulties in implementing franchising are the following:

- the possible collusion between the would be bidders;
- the existence of sunk costs which can hinder true competition;
- the establishment of contracts, which is costly and, of course, incomplete. This leads to many problems, in which opportunism can play a role.

Regarding infrastructure, we will pretend that franchising is not very efficient to regulate infrastructure management, but to make this point, we will need some developments on the infrastructure management activity itself and thus we will address this question in the fourth part of the report. Let us remark that the franchise for infrastructure are generally quite long and thus cannot lead to introduce strong competitive pressure. This is the case for the French highways, the underground of London, the future French Spanish high speed line Perpignan-Figueiras.

1.2.6 Yardstick competition

1.2.6.1 What we call “yardstick competition” ?

There is no precise definition of what is called yardstick competition, given that the associated theory has led to various ways of implementation, as described in chapter 3. However, we can distinguish two main senses given to the term “yardstick competition”.

- On the one hand, this expression refers to a regulatory framework, based on comparisons. It is a virtual form of competition between similar regulated firms, like Shleifer’s proposal (see next page). It consists in estimating what should be the best prices and subsidies, by comparing the performances of various regulated firms. The regulator, by setting the correct prices and subsidies, can lead the firms to produce an effort that increases the welfare.
- On the other hand, yardstick competition refers to the basic and relatively informal use of comparisons by a regulator who wants to improve its expertise and reduce the informational asymmetry he faces. In that sense, yardstick competition is an additional expertise tool used by the regulator to improve the efficiency of another regulatory framework (franchising, for example).

1.2.6.2 Yardstick versus benchmarking

Benchmarking is sometimes presented as a kind of yardstick competition. We do not consider benchmarking in this report, because it is, for us, undertaken by the firms, in order to improve their process or methods and not by the regulator.

Conclusions

Infrastructure provision has to be regulated, notably because it leads to a monopoly and thus market power. Others market failures are also present (opportunism, bounded rationality, and externalities). Infrastructure regulation is also problematic, mainly because of asymmetric information.

We will consider in this report that the aim of the regulator is cost efficiency.

Among the options offered to the regulators, we would like to stress the following points:

1. Government ownership seems desirable for an infrastructure separated from operation;
2. Price regulation is submitted to various objectives, but not primarily the cost efficiency of the infrastructure provider;
3. Rate of Return regulation should play a limited role, notably because of the Averch-Johnson effect;
4. Making the market contestable and franchising as well constitute options of limited scope because of the long term effect of many contracts and many maintenance operations. This point will be developed in the fourth chapter of the report.
5. Yardstick competition might be included in the regulatory framework of infrastructure management, but before going further, we will discuss the theoretical foundations of yardstick competition. This is the subject of the next chapter.

2. SURVEY OF THE MAIN THEORETICAL APPROACHES

In this part of the report, we detail yardstick competition mechanisms through a review of the main theoretical approaches. First, we present the principle of yardstick competition as defined by Shleifer and the ways of implementing it. In the two next sections we analyse the informational benefit of the comparisons in the two cases of informational constraints (moral hazard and adverse selection). Then we discuss the main limits to yardstick competition in a static context (correcting external heterogeneity) and in a dynamic one (investment incentives and collusion).

2.1. Principle of yardstick competition

In this section, we present generalities introducing to yardstick mechanisms. First, we discuss Shleifer's model, then we analyse how yardstick competition can be implemented.

2.1.1 The model of Shleifer

Shleifer (1985) is at the origin of theoretical mechanisms of yardstick competition. Its model, inspired by Medicare's observations (see chapter 3), defines the concept of yardstick competition relatively simply.

2.1.1.1 Model and assumptions

Shleifer considers $N \geq 2$ similar firms, operating on geographically separated but identical markets and producing the same output. He assumes that the firms face the same demand curve. Each firm i is characterised by its marginal cost c_i and its investment in cost reduction $R(c_i)$. For each firm i , the regulator sets the price p_i and distributes a subsidy (a lump-sum transfer), T_i . In order to set prices and subsidies according to the performances of the firms, the regulator compares each firm to its yardstick defined as described in the following paragraph.

2.1.1.2 The yardstick

For each firm i , consider:

$$\bar{c}_i = \frac{1}{N-1} \sum_{j \neq i} c_j$$
$$\bar{R}_i = \frac{1}{N-1} \sum_{j \neq i} R(c_j)$$

Each firm i is assigned its own "shadow firm" which serves as the benchmark in yardstick competition. Shleifer shows that the regulator can achieve the economic optimum by setting:

- the price of firm i , equal to the average marginal cost of the other firms: $p_i = \bar{c}_i$
- the lump-sum transfer to firm i , equal to the average investment in cost reduction of the other firms: $T_i = \bar{R}_i$

Using game theory, he shows that every firm's optimal strategy is to reveal its true cost c_i and investment in cost reduction $R(c_i)$. Thus each firm i is forced to compete with its yardstick, defined by the performances of the other firms. This scheme implicitly defines the costs that firm i has to reach and the investments in cost reduction, it should make, in order to have a positive profit.

2.1.2 *Implementation of yardstick competition*

Before reviewing applications of yardstick competition to utilities regulation (next chapter), we ask the three main theoretical questions related to the implementation of such a regulation framework. We also show, there are various ways of implementing yardstick competition.

2.1.2.1 *Correcting heterogeneity*

Through the simple model of Shleifer, we noticed it is highly necessary to correct the external heterogeneity, as, actually, the assumption of identical environment is not acceptable. The costs of the firms are also influenced by factors associated to each environment. Actually, Shleifer presents in the same article a simple model to correct external heterogeneity. As each yardstick is defined by the costs and investments of the other firms, uncorrected heterogeneity would bias the yardstick and then leads to an inefficient mechanism. This important question would be discussed hereafter.

2.1.2.2 *Yardstick competition and franchising*

Franchising, also called competition for the market, has been presented in chapter 1. It appears that the concepts of yardstick competition and competition for the market are clearly different. However, we pretend that those two ways of regulating firms are rather complementary.

On the one hand, competition for the market is efficient to the extent that the regulator has sufficiently information to prevent possible collusion between the bidders². It is possible to prevent collusion by reducing the benefit of it (which consists of a monopoly rent shared by the cartel members). That means that, if this profit can be sufficiently diminished, as it is possible, using yardstick competition, the regulated firms would have less incentives to collude. So yardstick competition can contribute to a successful implementation of competition for the market.

On the other hand, consider a firm – regulated through yardstick competition – which seems to be inefficient, according to comparisons. The inefficiency of that firm can be due to mismanagement or to the contract negotiated with the regulator. So it is possible to reduce this inefficiency by changing the management of the regulated firm; hence, the regulator should be able to manage the firm's exit of the market. However, given the specificity of transport infrastructure services, such an exit must not lead to service interruption. This is the reason why a recurrent bidding process (like franchising) seems to be an interesting solution. So competition for the market and yardstick competition are complementary and to a certain extent, enforced by each other.

2.1.2.3 *How to develop incentives ?*

Shleifer's model and other theoretical models use a financial mechanism linked to the results of the comparisons to enforce a competitive pressure. Those models show that gains expectation and penalty threat generate such a competitive behaviour. However, this is not the only way to reach this goal. In particular, the use of comparisons in addition to franchising induces competition between the regulated firms for the two following reasons:

- a reputation effect arises: no firm can afford to have a much worse image than the others. This effect appears as soon as the regulator widely spreads the results of its comparisons;
- a belief effect plays: if the regulator can convince its firms of its attachment to the results of the comparative mechanism, he can create an additional competitive pressure. In particular, the threat of non-renewal of the contract at the next bidding can induce such an effect.

² The collusion problem (the cooperative and, hence, uncompetitive behaviour between the bidders) is discussed more precisely in section 2.5.

So, yardstick competition can be implemented through different ways: in connection with franchising or not, with a financial mechanism or not. We now discuss in the two following sections some more sophisticated models which illustrate how the comparisons lead to reduce informational asymmetry.

2.2. Informational benefit of the comparisons in the moral hazard case

2.2.1 What is moral hazard ?

Moral hazard problem arises from the informational asymmetry, the regulator faces, concerning the regulated firm's behaviour. Because the regulator cannot control the firm's actions all the time, the regulated firm tends to adopt an opportunistic behaviour. This is the very reason why the regulator has to provide effort incentives for its regulated firm, as explained in chapter 1. However, the production of the firm depends not only on the effort it has exerted, but also on hazard (influence of a third party, of the weather, the overall economic situation...). Firms can face a commercial risk (if demand decreases) and an industrial one (if costs increase more than prices). Due to this hazard, the firm faces uncertainty which it can fear or not:

- consider a state-owned firm, operating a public utility service which suffers no interruption. This firm is quite sure that it cannot go bankrupt, because of its public ownership and the necessity of ensuring continuity of the service. It is ensured to be recapitalized sooner or later. Thus, such a firm does not fear uncertainty, knowing it is potentially ensured against risk. Such a behaviour toward risk is said risk-neutral.
- consider, now, a private firm. It cannot assume that its shareholders would accept every time the financial impact of random events. Hence, such a firm fears uncertainty. It is said risk-adverse. In a regulatory framework, the regulator has to pay such a firm a fixed incentive (a risk insurance), in order to offset the risk, the firm faces.

We now analyse, in this moral hazard case, how comparisons might reduce the uncertainty and then the risk insurance payment.

2.2.2 Modelling uncertainty

Many authors model the uncertainty which affects firm i 's production by the following elements:

- a common uncertainty parameter, η , which affects the whole firms of the sector. This variable includes mainly the situation of the sector (overall economic, social, political, competitive...).
- an independent particular risk, ε_i , which includes local impact on production (weather, difficulties caused by a third party...). The ε_j 's are assumed to be independent and normally distributed.

Thus, the risks ($\eta + \varepsilon_j$) they are facing are correlated to the extent that the regulated firms are similar. Mathematically that would be translated as :

$$\text{var}(\eta) \gg \text{var}(\varepsilon_j).$$

2.2.3 Theoretical results

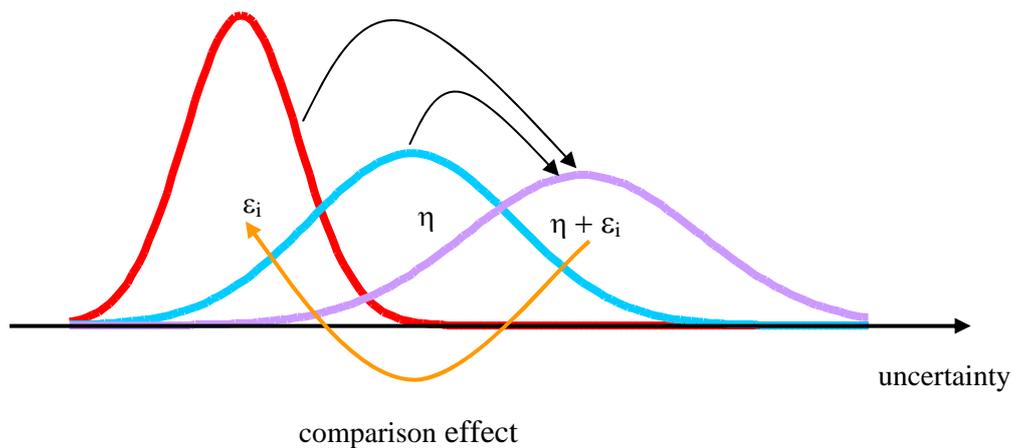
Holmström (1982) shows that, in such a configuration, the use of comparisons improves the welfare. The results are the following:

- if firms are risk-neutral, yardstick competition is as interesting for the regulator as the other regulation schemes which do not use comparisons.
- if firms are risk-adverse, the economic efficiency is improved by comparisons as soon as there is common uncertainty ($\eta \neq 0$). This improvement increases with the number of compared firms.

2.2.3.1 Reduction of uncertainty

The underlying idea is rather intuitive (see figure 2.1). Capturing all the relevant information about η , included within the outcome measures, leads the regulator to neutralise the impact of common uncertainty. The outcomes then depend only on the ε_j 's and the efficiency of the firms (assuming that external heterogeneity is corrected). Thus the risk, a firm faces, is reduced from $(\eta + \varepsilon_i)$ to ε_i .

Figure 2.1. **The comparisons reduce uncertainty**



Following that result, the higher the ratio $\text{var}(\eta) / \text{var}(\varepsilon_j)$ is, the more the comparisons are relevant. The regulator who reduces by this way the uncertainty can then decrease its risk-insurance payment to the regulated firms.

2.2.3.2 Application to transport infrastructure services

We believe that transport infrastructure services are rather few concerned by this mechanism. This is due to the relatively low uncertainty, the regulated firms face. $\text{Var}(\eta)$ is low because transport infrastructure service firms operate in a very stable environment: η takes mainly into account changes in the labour legislation and prices increase, which are not a major source of uncertainty. $\text{Var}(\varepsilon_j)$ is not so low, including mainly climatic events (which can cause severe and costly damages to infrastructures) and uncertainty related to the geological knowledge of the subsoil. However, it is relatively easy in transport infrastructure services to determine the financial impact only due to those random events. This enable the regulator to insure *ex post* its regulated firm against risk.

All the more, transport infrastructure service firms are rather risk-neutral, because of their public ownership and/or social utility (the service suffers no interruption). Therefore, although yardstick competition cannot have a negative effect on the regulation of transport infrastructure services, its implementation in such a moral hazard case does not seem very interesting. However, regulators of infrastructure companies could be rather concerned by the adverse selection situation which is addressed in the following section.

2.3. Informational benefit of the comparisons in the adverse selection case

2.3.1 What is adverse selection ?

Adverse selection problem arises from the informational asymmetry, the regulator faces, concerning the regulated firm's identity (its productivity). For a given production and a given payment by the regulator, consider two types of firms:

- a low-efficient one, which needs to exert a lot of efforts to provide the service required. Its income (relating to its efforts) are just sufficient to ensure its participation.
- a high-efficient one, which needs to exert only little efforts to provide the same service. Its income (relating to its efforts) ensure it a quiet life. However, such a firm could provide a much better service, exerting much more efforts; but it would do so only if it receives sufficiently high incentives. This transfer is also called informational rent.

In the adverse selection problem, the regulator cannot distinguish (because of informational asymmetry) the high-efficient firm from the low-efficient one. Then it cannot costlessly force high-efficient firms to exert a lot of efforts, rather than to have a quiet life. To prevent those firms from passing themselves off as low-efficient firms, the regulator has to:

- encourage them to reveal their high-efficient identity; to reach this goal, it has to pay them the informational rent;
- discourage them to adopt a low-efficient identity; to achieve this objective, it has to reduce the interest of low-efficient firms' situation (by decreasing the level of service required, and then their income); this production distortion leads the regulator to reduce the informational rent that it has to pay.

Thus, in order to optimise the welfare, the regulator has to propose the firms a revealing menu of contracts. This menu should include a contract (low production and low income) chosen by the low-efficient firms and another contract (high production and high income) for the high-efficient firms. We now analyse, in this adverse selection case, how comparisons might reduce the informational asymmetry on firms' identity, and hence the informational rent and the production distortions.

2.3.2 Modelling productivity

The firms' productivity model is designed like uncertainty model: one distinguishes a common element (whose impact would be reduced by the comparisons) and a residual particular one. So, many authors model the productivity parameter, β^i , which characterises each firm i (assumed to be risk-neutral) by the following elements:

- a common part, $b \in \{b_1, b_2\}$, which is the same for every similar firm of the market. This variable reflects statutory, regulation, organisational... particularities of the market. b_1 corresponds to high-efficient firms, whereas b_2 refers to low-efficient ones.
- a particular part, ε^i , which is assumed to be independent and identically distributed in a given interval. This individual characteristic includes the effects on the efficiency of firm i 's decisions (investments, management...).

2.3.3 Theoretical results

Auriol (2000) proves that, in such a configuration, the use of comparisons improves the regulation efficiency (for more details, see annexe). She considers that the regulator proposes the firms a revealing menu of contracts. Using game theory, Auriol shows that firms are encouraged to choose the contract which corresponds to their productivity parameter, β^i . Hence, they reveal their common part, b_1 or b_2 . Knowing this information, the regulator is able both to reduce the informational rent of the high-efficient firms (b_1), and to suppress the distortions on the production of low-efficient firms (b_2).

2.3.3.1 *Reduction of informational rent*

Following the above mechanism, the regulator can detect inconsistent announcements (when a firm announces a characteristic corresponding to a common part $b = b_2$, and the others announce characteristics corresponding to $b = b_1$). So, the regulator can retaliate against the firm of such an inconsistent announcement (excluding it from the market, for example), in order to make the mechanism revealing. So, a high-efficient firm cannot behave anymore as if it were low-efficient. The informational rent is then reduced.

2.3.3.2 *Suppression of production distortions*

Given that b_1 firms cannot pretend anymore to be of type b_2 , it is not necessary for the regulator to maintain the distortions on b_2 firms' production. Thus, the comparisons lead the regulator to make those firms exert the optimal effort level.

2.4. **Correcting external heterogeneity**

After having presented Shleifer's model, we noticed the need to control for the external heterogeneity which influences firm's performances, and then the yardstick they are compared to. In this section, we illustrate how external heterogeneity impacts on the firm's performances and how it can be corrected. One usually distinguishes the two following types of heterogeneity:

2.4.1 ***Endogenous heterogeneity***

This section refers to Bivand & Szymanski (1997) who analysed this effect of endogenous heterogeneity, that they define as "spatial dependence effect".

2.4.1.1 *What is endogenous heterogeneity*

For the situation under review, the transport infrastructure regulation, endogenous heterogeneity arises as soon as the regulatory framework differs between the compared firms. If yardstick competition is implemented by a unique, centralised regulator, every regulated firm faces the same requirements. In such a case, there is no endogenous heterogeneity. However, if there is decentralisation, each firm faces its own regulator's requirements. Endogenous heterogeneity may arise from particular policies of each regulator.

For example, a regulator (1) could require a very high quality of service from its firm 1 and pay the (relatively high) associated price. In this case, firm 1 seems inefficient (because of its high costs) if the service quality is not taken into account. Another regulator (2) could be financially benevolent toward its firm 2. This firm would appear relatively inefficient, for its costs are high, due to the benevolence of its regulator. If comparisons done by a third regulator (3) cannot take into account those particular policies, they would be biased in favour of its firm 3, which would be compared to apparently inefficient firms.

Such policies generate externalities which influence the firms' performances and, subsequently the comparison yardstick.

2.4.1.2 *The case of transport infrastructure services – how to correct endogenous heterogeneity*

Concerning transport infrastructure services, regulation is rarely centralised: on the one hand, local - urban and rail transport - services are often regulated by local authorities. On the other hand, main road infrastructures (highways, civil engineering works) are nationally regulated, but apparently loosely. In this context of decentralised regulation, the different regulators' involvement in the development of their transport infrastructure services would be a major source of endogenous heterogeneity. Contrary to other network industries (water and energy supply, telecommunication)

which have to serve the whole population, always and everywhere, transport infrastructure services are frequently defined by the local or national government and the decisions are motivated by political considerations.

Those decisions impact costs of service and the level of traffic.

- To reduce the impact on costs, it is highly necessary to correct this heterogeneity by including the corresponding factors (quality and/or frequency of service...) required in the estimated cost function, as it should be done to correct exogenous heterogeneity (see above).
- To reduce the influence on the traffic level, the regulator has to base its comparisons on costs which are independent from the level of traffic³. For example, a urban transport regulator should compare average costs per bus.kilometer rather than average cost per passenger.kilometer, which depends on the firm's performance, but also on the fare, the frequency of services... set by the regulator.

2.4.2 Exogenous heterogeneity

For transport infrastructure, exogenous heterogeneity is caused by the environment. That kind of heterogeneity is due neither to the firm nor to the regulator. This environmental heterogeneity arises from geographical constraints (relief, flood risk, network spatial structure...), demographic differences (density of population), social particularities (delinquency rate) and many others...

In order to correct those sources of distortion, the solution consists in introducing a measure of those factors in the estimation of the cost function. It means that the accounting cost of the service has to be regressed through an econometric process before being used in the comparisons. But this way of doing leads to other difficulties, that we expose briefly in the next section.

2.4.3 Choice of a cost function

We have just seen that correcting heterogeneity needed to specify a cost function including the different factors of heterogeneity. However, this solution should be reserved to the comparison of sufficiently similar firms, for several reasons.

- First, it is not easy to include every source of heterogeneity. In particular, the quantitative measure of some of them may be very difficult.
- Secondly, the mathematical form chosen for the cost function influences the results of the correction.

Thus, the choice of the cost function (specification and form) may impact the heterogeneity correction, which would not be perfect. This is the very reason why yardstick competition should be applied to similar firms, operating in relatively close environments which do not need a wide heterogeneity correction.

However, concerning infrastructure transport services, the financial impact of the main factors is relatively well known. For example, the maintenance cost of an infrastructure is predictable, given traffic characteristics (number of vehicles, speed, weight, etc.). Hence, it is rather easy to ensure a good correction of the heterogeneity between compared infrastructure transport services.

2.5. Some theoretical limits

Before concluding this chapter, we analyse yardstick competition in a dynamic context, since, actually, the robustness of a regulation framework has to be considered through the time. We discuss

³ Nash (2000) developed such an analyse concerning railways : "Governments [...] frequently intervene in the pricing and output decisions of railways. Performance measures for these railways then typically provide information on a mixture of the performance of the management and of the institutional setting in which it operates. For passenger services it is not uncommon for governments to effectively control the timetable as far as the frequency of service on each route [...]. In this situation, arguably the government becomes the customer, and the output the railway produces is a certain level of service, rather than transport for a number of people."

two common limits of a regulation framework: the investment incentives, which have to reach an adequate level (neither under-investment, nor over-investment), and the collusion, which we have already mentioned concerning franchising.

2.5.1 Reaching an adequate investment level

2.5.1.1 Spillovers and under-investment

Dalen (1998) shows that yardstick competition could reduce investment incentives in some cases. Taking Auriol's model framework, he considers that every investment can be split in two parts:

- a "firm-specific" part, which increases only the efficiency of the investing firm. That is to say that this investment allows the firm to improve its own productivity (ϵ^i),
- an "industry-specific" part, which increases by externality ("spillovers") the common efficiency of all the firms of the sector (b).

Of course, yardstick competition promotes firm-specific investments. On the contrary, when investments benefit to the other firms, yardstick competition reduces investment incentives. This can be particularly the case of research and development, whose outcomes may spill over into the other firms at a very low cost, compared to what the first firm has invested.

So it is necessary, before implementing yardstick competition, to evaluate its impact on the research and development activities in the sector. However we believe that this problem of spillovers and under-investment should not particularly affect transport infrastructure services.

2.5.1.2 Regular long term investments

In the first chapter, we explained that difficulties could arise because private firms are rather short term oriented. This leads to foresee another theoretical investment problem when yardstick competition is applied. There is a risk that an excessive competitive pressure on the current expenditure of a regulated firm leads it to delay or reduce long term investments. This problem which affects particularly infrastructure services is due to the difficulties of long term regulation. We would discuss it in the last chapter.

2.5.2 The collusion between compared firms

2.5.2.1 What is collusion ?

Collusion is usually defined as the co-operative behaviour between firms which should – on the contrary – compete against each other. Such firms can agree to form a cartel and to co-ordinate their prices or informations, in order to reduce the competitiveness of the oligopolistic market. This leads them to benefit from higher (monopolistic) prices and to increase their profit at the expense of the consumers. Hence, the resulting allocation of resources is sub-optimal. This the reason why collusion has to be avoided.

Note that a collusive agreement requires the participation of every firms operating on the associated market. If one does not collude, it can practice lower prices than the collusive ones, and hence it gets the whole market and the cartel fails. Thus, it is difficult for the colluding firms to preserve the cartel from any deviation, given that each of them has opportunistic incentives to deviate.

Moreover, collusion between firms might occur even in a non co-operative context (see Tirole, 1993). This tacit collusion arises from the threat of a vigorous price war, after one firm cuts prices. This provides strong incentives, not to compete actively so that a relatively quiet life can be maintained on the market.

2.5.2.2 *The reasons for which collusion is likely to appear*

A collusion threat arises when firms are regulated by yardstick competition because they foresee, they get zero rent from the mechanism proposed by the regulator if they play non co-operatively. Hence, they are willing to co-ordinate their messages to countervail the regulator's power.

In particular, the "revealing principle" used by Auriol may become inefficient if firms collude. Consider the model of the firms' productivity in the case of adverse selection, with a parameter $b \in \{b_1, b_2\}$, distinguishing high-efficient firms from low-efficient ones. At first sight, high-efficient firms seem to be able to collude and announce a low-efficient characteristic. This would permit them to exert a lower effort than the optimal one. In this case, collusion, directly distorting the comparison yardstick, reduces the incentive effect of yardstick competition. However, Auriol's particular model is so designed that the mechanism suppresses any collusion incentive. The best choice for the firms, whatever their productivity characteristic is, consists in choosing the contract corresponding to their productivity.

2.5.2.3 *Why the risk of collusion is not always so high*

The theoretical threat of collusion has not been observed in practice in the sectors described above. Many reasons can be proposed:

- The high number of compared firms is an important limit to the development of collusion. This is due to the poor robustness of collusive agreements when the number of participants increases, because it becomes difficult to involve each of them in the cartel and to prevent any deviation.
- The comparisons done by the regulator lead it to detect easily any deviation from a competitive behaviour towards a collusive one.
- Moreover, comparisons based on accounting data (expenses, investments...) reduce the ability of the regulated firm to hide or modify their information.

For these reasons, we believe that collusion is not so an important risk, when firms are regulated by yardstick competition.

Conclusions on the main theoretical approaches of yardstick competition

To conclude this chapter, we can note that yardstick competition seems to be an interesting way of regulating similar monopolies. The use of comparisons provides strategic information for the regulator. In particular, it leads to reduce the informational asymmetry, the regulator faces. Some mechanisms might hinder or at least limit collusion. The theory is flexible enough to give rise to various kind of applications. That is what we will observe in the next chapter.

3. APPLICATIONS OF YARDSTICK COMPETITION TO UTILITIES REGULATION

In this chapter, we present the main applications of yardstick competition to the regulation of various utilities (hospitals, water utilities...), with special attention to transport services (Japan railways industry and Norwegian buses). The objective of this chapter is to show the many different ways of implementing a regulation scheme related to some kind of yardstick competition.

3.1. The activity based reimbursement of hospitals:

Historically, the reimbursement of hospitals is the first application of yardstick competition. Medicare, the American program, began in 1983, two years before Shleifer's theoretical proposal. At the beginning, this regulation scheme was very close to pure yardstick competition.

3.1.1 Activity-based reimbursement and yardstick competition

3.1.1.1 Principles

Given the inefficiencies of hospital systems in the OECD countries (high costs or long queues), many countries have reformed it, using activity-based financing⁴. However, both the objectives and the implementation of this financing system vary a lot between the different countries. The common point of those activity-based systems is the definition of so called "diagnosis related groups" (hereafter DRG). A DRG is a group of illnesses or pathologies, which needs the same hospital treatment. Hence, it is possible to define and to compare the costs of each hospital for every DRG.

3.1.1.2 Objectives

The three main objectives of an activity-based healthcare system are the following:

- to fit the financing to the activity,
- to encourage the hospitals to reduce the costs increase,
- to promote competition in order to reduce some rents.

The top objective of each government depends on the key inefficiency observed. On the one hand, countries facing too high costs (US, Germany, France) would be interested by the costs growth reduction. On the other hand, countries facing too long queuing (United Kingdom, Scandinavia) would prefer promoting competition in order to encourage hospitals to improve their productivity.

3.1.2 Implementation and effects

It appears that the implementation of an activity-based system depends on the top objective followed by the government.

Governments aiming at reducing the costs of their hospital system use price regulation, similar to Shleifer's proposal. However, the hospital reimbursement systems are rather mixed: they are not uniquely related to the comparisons. For example, Medicare, the American health insurance for old and handicapped people, is based on such comparisons. Since 1983, DRG have been defined and, for each of them, the reimbursement price is based on the average of the costs (corresponding to this DRG) observed in the hospitals. Of course, external heterogeneity is corrected by including in the comparisons demographic and geographical data.

Although it is difficult to evaluate the efficiency of a price regulation, it seems that the effects of activity-based reimbursement are positive: hospitalisation length decreased, costs growth was reduced, whereas the healthcare quality was preserved.

The other governments, whose objectives are to reduce hospital queues, promote competition through comparisons. For a few years, the British government has compared the costs of its hospitals for each DRG. However, it seems that the competitive pressure related to the use and spread of the comparisons was not strong enough in order to achieve productivity increases for hospital.

⁴ See DREES (2002).

3.2. Water utilities in Great Britain

This case study is interesting, because it illustrates the benefits of the relatively informal use of the comparisons by a regulator.

3.2.1 *The regulatory framework and the comparisons*

The 26 Welsh and English water utilities are regulated by the Office of Water Services (OFWAT). Although firms compete for the market, the regulator has to exert a pressure on them, given the very long duration of the contracts (25 years). Prices are regulated through a price cap mechanism (see chapter 1). The initial level of the price limit and its evolution (the anticipated productivity gains) are set by the OFWAT every five years. Moreover, the regulator compares the firms' performances every year, in order to promote "comparative competition". Performance measures are based on service quality, accounting and financial data and are used as exposed next section.

3.2.2 *How are used the outcomes of the comparisons?*

As explained in the previous chapter, such an informal use of the comparisons leads to a competitive effect, because of a reputation effect, and to an increase of the regulator's expertise.

3.2.2.1 *A reputation effect*

The comparisons lead to a reputation effect which encourages the firms to behave competitively, because no one can afford to have a much worse image than the others. This effect arises because of the public spread of the outcomes of the comparisons. The OFWAT gives in a yearly report the results of the firms, concerning their cost efficiency. Moreover, the report includes tables of the data related to service quality, so that customers can estimate the performances of their water utility. This is proved to exert a competitive pressure on the firms.

3.2.2.2 *An increase of the regulator's expertise*

Comparisons lead also the regulator to increase its expertise concerning two points. First, the OFWAT uses the comparisons to set the parameters of the price cap scheme. Price ceiling and its evolution are defined for each firm given the outcomes of the comparisons, among other data. Secondly, the comparisons lead the regulator to detect some particular inefficiencies (for instance, too many leaks in a given network) and then to reduce the informational asymmetry, he faced at the beginning.

It has been shown⁵ that the improvement of the firms' efficiency was more important for the firms which were inefficient at the beginning of the regulatory period. This proves that comparisons have been playing a role – in addition to the price cap – within the competitive pressure exerted by the OFWAT on the firms.

3.3. Dual-sourcing

Dual-sourcing consists in introducing a second firm on a monopolistic market. It sometimes appears that the advantages of a duopoly structure can be preferred to the duplication of fixed costs, given the difficulties to regulate a monopoly. This is particularly the case when the fixed costs are due to large investments in research and development. The two following examples show that dual-sourcing offers to the project manager a mixture of co-operation and competition with various advantages.

⁵ See OFWAT (1997).

3.3.1 US Defence Department

The American Defence Department has been using dual-sourcing since 1980, especially for important expensive systems⁶. For instance, the Advanced Medium Range Air-to-Air Missile, the Tomahawk cruise missile, fighter aircraft engines F15 were produced by two competing firms, following the dual-sourcing approach.

The process is the following: after competition at the design stage, one firm is selected by the Department of Defence to develop its proposal and to undertake the initial production. After initial production, the Department of Defence can transfer technology to a second firm and force the two potential producers to compete for a production contract. Technology is transferred through a “learning buy” in which the second firm receives generous compensation for delivering a few products.

The benefits of this dual-sourcing practice are the following.

- The reduction of the expected marginal cost due to the use of yardstick competition between the two producers.
- Avoiding the capture of the Department of Defence by a sole firm, having a large informational and learning advantage on its potential competitors. One of the objectives of dual-sourcing is to share knowledge, so that competition can occur for further development and/or production contracts.

3.3.2 The Petronas Twin Towers of Kuala-Lumpur

The Petronas Towers of Kuala-Lumpur are the tallest building in the world. Each of the twin towers was built by a different firm: Samsung built the first one, whereas Mitsubishi constructed the other.

The project manager decided to use dual-sourcing for the following reasons:

- the technical difficulties to build so tall towers led to foresee some uncertainty (problems that would occur and would have to be solved quickly). So dual-sourcing allowed both contractors to learn from each other if one of them faced a difficulty.
- having two separate construction teams has developed a healthy competition between both parties, to see who would do a better work. For instance, Mitsubishi who had begun the second tower one month after Samsung finished the construction at the same time. In that sense, dual-sourcing has provided incentives to compete.

3.4. Transportation services

The two following cases of regulation frameworks related to yardstick competition are interesting for two main reasons. First, because they concern transport services (Japanese railways and Norwegian buses). Secondly, because the use of comparison is not informal as in the two previous cases. Conversely, comparisons are used in the financial regulation mechanism. This is why we develop more precisely the underlying yardstick mechanism for both cases.

3.4.1 Yardstick competition implemented in Japan railways industry

This part of the report owes very much to a recent article (Masaru Okabe, 2004) and would not have been possible without the help of Makoto Ito and Fumitoshi Mizutani. The yardstick competition system in Japan is original and seemingly at least partly successful (Mizutani, 1997). The railways companies submitted to yardstick competition are the following :

- the 15 major private companies;

⁶ See Riordan and Sappington (1989).

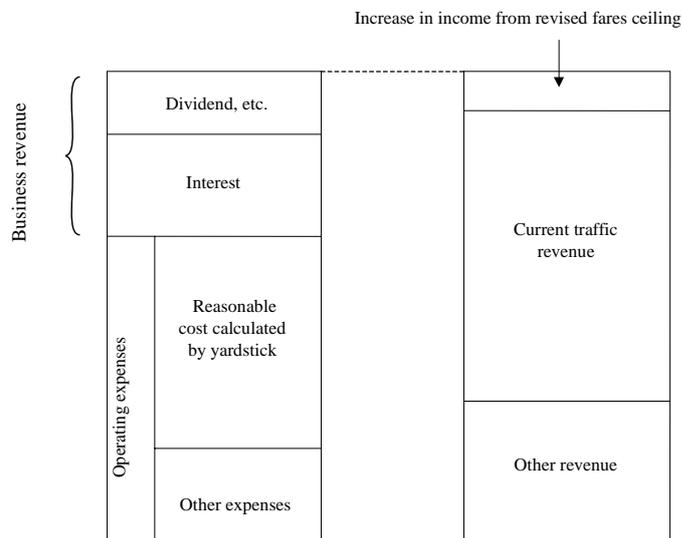
- the 6 Japan Railways (the JRs, the six regional companies created in 1987 when JNR was broken up into six passengers companies bound to be privatised and a freight company);
- 10 public subways companies.

With those 31 companies, yardstick competition is now applied to railways carrying about 95 % of rail passengers in Japan. We present successively a general overview of the system, the yardstick competition mechanism and a brief summary of the assessment made by Mizutani.

3.4.1.1 General overview of the incentive scheme

The regulatory scheme is a combination of price cap, ROR and yardstick competition. The fares must be under a ceiling (they can be significantly under this ceiling). This ceiling is given by a calculation based on a “reasonable cost”, which is the result of yardstick competition. Thus, a comparatively cost efficient company can have a greater “business revenue”. But there is a trade-off between revenue and lower fares, as this procedure allows company to change fares, under the ceiling, by simple notification to the government. However, to avoid unfair parallel competition with cross subsidisation, some fare reductions are submitted to government approval.

Figure 3.1 Overview of the incentive scheme in Japanese railways



Source : adapted from Okabe, 2004.

The reasonable cost is calculated by a yardstick competition process, which will be describe below, the total cost is determined by ratios. The procedure depends on the kind of firm regulated, but this implies a fair profit on capital invested. The ceiling is given by the total cost, including the reasonable cost calculated by the yardstick.

3.4.1.2 The yardstick competition

There are five costs categories (see table 3.1). For each of those costs, the unit cost is calculated by dividing the costs by the number of facilities (associated volume).

Table 3.1. Facilities and explanatory variables for each kind of costs

costs	Facilities (associated volumes)	Main Explanatory variables
Tracks	Track length	Rolling stock km per track km (Log)
Catenary	Catenary length	Electric multiple unit km per catenary km
Rolling stock	Number of rolling stock	Rolling stock km per rolling stock
Train operation	Route-kilometer	Train km per route km (Log)
Stations	Number of stations	Number of passengers per station.

Source: Okabe, 2004.

Then for each category of cost, a model is calibrated with some explanatory variables supposed to be correlated with the unit cost. The main explanatory variables are given in table 3.1, but the model might differ according to the railways class. For track costs (mainly maintenance), the details are given in table 3.2.

Table 3.2. Calculation of the standard unit cost for maintenance cost (maintenance costs per kilometer of line)

Group of railways	Regressors
15 large privates rail companies	Tunnel and bridges percentage Log(traffic density)
6 JR rail companies	Log(traffic density) Amount of snowfall
10 public subway companies	Log (traffic density)

With those models, it is possible to calculate, for each category of costs and for each railways company, the base unit cost, which might be interpreted as the unit cost that the railway should have if this company has the average cost as calculated by the model.

Then the standard cost, for each category of cost, is calculated for each company by multiplying the base unit cost by the actual associated volume.

Finally, the total of the “yardstick” costs can be calculated. This cost is called the reasonable cost.

Then there are two possibilities.

If the actual cost is inferior to the reasonable cost, half of the difference is added to the actual cost and this becomes the new reasonable cost. This corresponds to sharing the benefit of higher productivity between the consumer and the producer.

If the actual cost is bigger than the reasonable cost, the reasonable cost is not changed. That means that the company will have to increase productivity in order to reach equilibrium.

There is another incentive mechanism. If a company decreases its productivity, as measured by the difference between actual cost and reasonable cost, the reasonable cost is diminished by half of this

difference. To the contrary, in case of cost efficiency improvement, the reasonable cost is increased by half of the improvement.

The total cost and the ceiling are calculated as previously indicated.

3.4.1.3 The effect of yardstick competition in Japan

The regulation system described here is the result of a reform which took place in 1997. The only published analysis in academic journals is the one of Mizutani (1997). It was therefore carried out before the reform. Nonetheless, some points are noteworthy :

- large private operators submitted to yardstick competition have made some significant effort to improve cost efficiency;
- the effect of yardstick competition is not only to reduce operating cost but also to reduce the variance of operating costs, as would have been a “true” competitive effect;
- however, measured as difference with a reference model, the managerial efficiency improved for only 40 % of operators.

It would be very interesting to have new measures of the effect of yardstick competition under the new regime, particularly including the six JRs.

3.4.2 Norwegian buses

This section owes very much to Dalen and Gomez-Lobo’s recent paper (2003). In Norway, responsibility for local bus transport is decentralised to the 19 regional governments (counties). Each county is free to choose its own regulatory policy; many of them have implemented for years yardstick competition. Yardstick competition is used by some regional regulators to determine the level of subsidy granted to the companies (there is a large number of bus companies in each county).

The level of subsidy granted to the n^{th} company, S_n , is equal to the difference between expected traffic revenue (R_n) and expected costs (C_n):

$$S_n = R_n - C_n = P_n y_n^e q_n - \sum_k c_{k,n} - c_{\text{adm},n}$$

- concerning the expression of the expected revenue, P_n is the average fare level set by the regulator for the bus services of company n , y_n^e is the expected number of passengers per kilometre produced and q_n is the total number of kilometres produced, defined by the regulator.
- concerning the expected costs, C_n is the sum of the k -inputs (drivers, fuel, maintenance, bus capital) costs, the $c_{k,n}$, and the administration costs, $c_{\text{adm},n}$, set proportional to $\sum_{k \neq \text{fuel}} c_{k,n}$.

Each input cost $c_{k,n}$ of company n is defined as follows: $c_{k,n} = p_k \sum_i \sum_j a_i^j q_{i,n}^j$, where p_k is the yardstick unit (per kilometre produced) price for input k , $\sum_i \sum_j a_i^j q_{i,n}^j$ is the sum of the kilometres produced by company n , weighted by the route type, i , and the bus type, j , (coefficients a_i^j). This weighted sum corresponds to the correction of heterogeneity due to route type and bus type.

So, yardstick competition plays because the subsidy granted to a company is related to the benchmark prices of the k -inputs, the firm should buy to produce its transportation service.

Dalen and Gomez-Lobo (2003) proves that bus companies regulated under yardstick competition become more efficient and reduce their costs faster than the others; in this case, yardstick competition provides more dynamic incentives for cost reduction by operators.

Conclusions on the applications of yardstick competition to the utilities regulation

Through various case studies, we have illustrated how regulation schemes could be more or less related to yardstick competition. This proves an interesting property of yardstick mechanisms: their flexibility, in that sense that they can be used in multiple cases, given various constraints and objectives. Moreover, yardstick competition seems particularly interesting for some kind of transport services regulation. We now discuss the particular case of transport infrastructure services.

4. YARDSTICK COMPETITION IN INFRASTRUCTURE PROVISION

In this final chapter, we will address the question of implementing yardstick competition for infrastructure provision. To that end we will first analyze the activity of infrastructure provision. We will see that it consists of a set of different activities. We will then consider step by step the different activities undertaken by an infrastructure provider. This part of the report concerns principally railways, but some points will be relevant for tolled motorways.

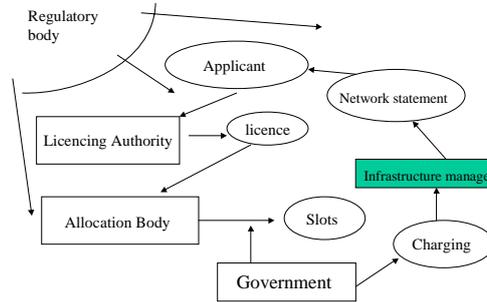
4.1. Analysis of infrastructure provider activities

Even if infrastructure is frequently owned by the state, that does not imply that the infrastructure manager activities should stay away of any competitive pressure. Before going through the possibility to implement yardstick competition at various stages of the activities of the infrastructure manager, it is useful to analyze more precisely what are the tasks of an infrastructure manager. This will allow us to define what we call an infrastructure manager.

Basically, infrastructure management consists in providing transport capacity at some prices and at some costs. To provide capacity, infrastructure has to be build, maintained and « operated ». Maintenance work can be further divided into light maintenance and renewal. All those operations have to be undertaken under the supervision of what we will call an infrastructure manager. But, actually, the infrastructure manager activities might encompass various tasks. Notably, the initial investment policy might not belong to the infrastructure manager. Because the government might take the responsibility to set up new infrastructures, contribute to the investment and then give the responsibility of operation to the infrastructure manager. Furthermore maintenance tasks may be carried out by the infrastructure manager or not.

Moreover, regarding railways, the infrastructure operations are parts of a complex system. As derived from the European directives the access system should be as described in figure 4.1.

Figure 4.1. **The institutional context of European Infrastructure manager (railways)**



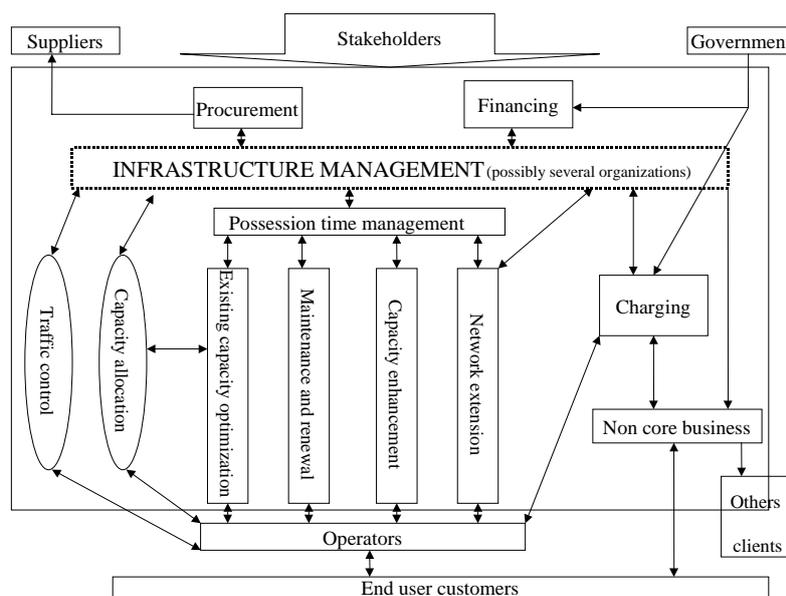
It follows that a certain number of activities might or might not be under the only responsibility of the infrastructure manager: notably slots allocation and infrastructure charging. Thus the final traffic on the line might not be the result of the infrastructure manager performance. Hence, the actual traffic on the infrastructure is to be considered as an exogenous factor. It follows also that any assessment of the relevance of the pricing system should be led independently of the regulation of infrastructure manager activities.

- To a certain extent we can say that infrastructure managers have a multiple production function:
- the first stage consists in the production of infrastructure, the output being the infrastructure by itself, considered as a physical object described by its material features;
 - the second stage consists of the production of capacity, the quantity and quality of slots that might be offered on the infrastructure;
 - the third one leads to the final “product”: the production of traffic.

This concerns only railways. For highways, only the first stage is relevant, and the final product is the production of capacity, partly the result of the initial design of the highways, and thus not fully under the responsibility of the infrastructure manager.

Regarding railways, the links between the different tasks and the various stakeholders can be represented with a “core process map” (figure 4.2).

Figure 4.2. **Core process map of “infrastructure management”**



Source: Improverail research project for the European Commission

The possible existence of a slot allocation body and of a traffic control organization, independent of the infrastructure manager as well as the fact that charging might be decided by the state with various objectives, make the central processes of the infrastructure manager the following:

- existing capacity optimization;
- maintenance and renewal;
- capacity enhancement;
- network extension.

All those processes have a link with the so called possession time, the time necessary for maintenance. We will distinguish maintenance from other operations, because maintenance costs can be easily observable.

So, in order to see to what extent yardstick competition can be implemented, we will consider here successively the three steps: infrastructure initial construction; maintenance; and other infrastructure operations. Finally, we will add a few remarks on the “production of traffic”, because many railways in the world are integrated. We will not discuss here the question of vertical separation, which is out of the scope of this report. Nonetheless, we can observe that vertical separation make yardstick competition for maintenance easier, because the necessary data are more easily observable. Nonetheless we will consider implementing yardstick competition for infrastructure with integrated firms.

4.2. Initial construction

Regarding initial construction, the same analysis is valid for road infrastructure and rail infrastructure. The usual way to introduce competitive pressure on the suppliers is competitive tendering. Could yardstick competition be jointly used with competitive tendering ?

The main difficulty with yardstick competition is that there is a very large exogenous heterogeneity of civil engineering works (engineering structures).

A kind of yardstick competition might be useful for simple civil engineering works such as building an infrastructure (road or rail) in a flat countryside. We will not expand on this, given the fact that yardstick competition might have only a limited role to play in putting under pressure the suppliers of initial construction. May be what is called “dual sourcing” could be usefully implemented, if the infrastructure has two comparable parts.

4.3. Maintenance

Regarding the maintenance and operation of infrastructure, we have to distinguish the case of road infrastructure and rail infrastructure. Concerning rail infrastructure the European Union is to be considered under a different point of view, because for those countries, infrastructure management is to be separated from train operations. Actually, in the European Union, the infrastructure manager is not providing a final product and the customers are the railways companies (the “railways undertakings” in the official terminology of the European institutions) who operate trains. It follows that for railways infrastructure, maintenance can be:

- subcontracted to independent companies;
- undertaken by the main railways company (the “historic operator” in the European context);
- undertaken by the infrastructure manager;
- a combination of the above possibilities.

For the countries where the railways are vertically integrated (the US, Japan, ...), the only choice is to rely more or less on outsourcing for maintenance.

During the first years following the British Railways reform, maintenance was largely undertaken by independent maintenance companies. The tendency is now that the infrastructure manager (Network Rail, formerly Railtrack) tends to reintegrate the maintenance. In France, the maintenance of

the track is undertaken by the historic operator, under the responsibility of the infrastructure manager. So far, this historic operator is the only licensed French railways undertaking.

In the case of outsourcing of maintenance works, it seems that there is a clear possibility to implement yardstick competition, the infrastructure manager being the authority and the maintenance companies being regulated. So, we have two possibilities:

- yardstick competition for the infrastructure manager's suppliers,
- yardstick competition on the infrastructure managers.

This corresponds to different levels:

- the cost of elementary maintenance operations,
- the relevance and the cost efficiency of the maintenance policy.

We will not expand on the first level: yardstick competition might be used jointly with competitive tendering, in order to minimise the cost of external works and in order to compare it to works done by the infrastructure manager itself or by the historic operator. Regarding the question of the possible integration of maintenance companies, we are facing the trade-off between the benefits of integration and the benefits of competition. The transaction costs argue, in principle, for integration, but the price to pay is that there is no more competitive pressure.

Regarding the second level, which consists in the comparison of maintenance cost, the question is, in short, how to judge that a maintenance policy is appropriate ? Several remarks can be offered.

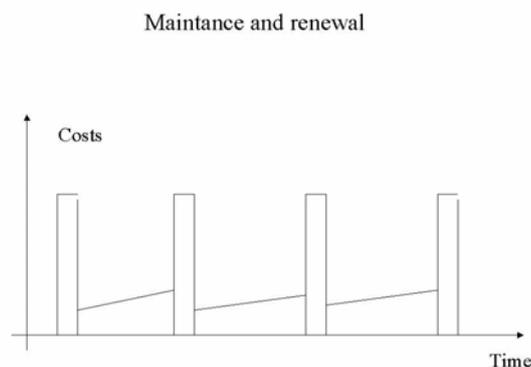
1. Maintenance has to be suited to the kind of traffic on every segment of lines. It follows that the set of lines under the responsibility of the infrastructure manager has to be broken down in different categories, notably according to the number of ton transported per year.

2. Maintenance consists of two kinds of works:

- light maintenance;
- renewal, which consists in heavy works undertaken to get a quasi new infrastructure.

So, ideally, the maintenance policy has to be compared over the life cycle cost, which includes Maintenance and Renewal (see fig 4.3)

Figure 4.3. **Simplified representation of the life cycle costs**



3. Thirdly, light maintenance is carried out during a time, hereafter called the possession time, during which no train can run. For some obvious reasons, the maintenance costs decrease when the possession time increase. The fact is that the opportunity cost of the possession time, tends to vary a lot, according to the traffic on the line and notably according to the degree of congestion on this line.

If the line is operated with sufficient free capacity, the opportunity cost of the time devoted to maintenance is zero.

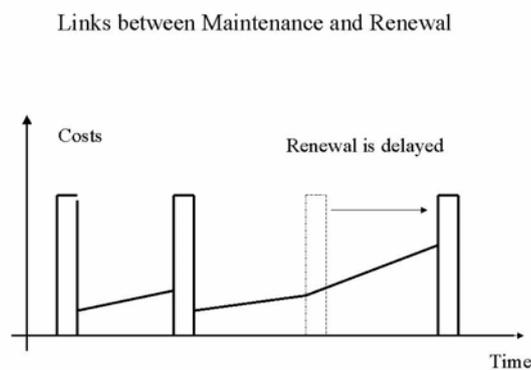
The last remark establishes a link between maintenance and what we call infrastructure operation. That means that for heavy loaded lines the assessment of the relevance of the maintenance policy cannot be isolated from the general management of the traffic.

What are the conclusions of the previous remarks concerning yardstick competition for maintenance ?

First, the traffic, probably measured in tons, has to be taken into account and thus included in the external factors which might affect the maintenance cost. Whether the dependency is purely linear or more complex has to be studied.

Secondly, any yardstick comparison should take into account the life cycle cost and not only the current maintenance cost. This is not an easy task. A first solution could be to have a dummy for light maintenance and renewal and to compare uniquely what is comparable. But those two kinds of costs are linked. If the infrastructure manager delays the renewal during a significant time, the maintenance cost increases. (fig 4.4)

Figure 4.4. **The absence of renewal increases the maintenance costs**



Another solution, more complex and needing more data would be to take the cost, discounted with the classical methods, on the entire life cycle costs.

If the rail networks to be compared are big enough and if there are a sufficient number of networks taking part in the yardstick competition, the problem of life cycle costs might be less disturbing. For each network, during an average year, some parts of the network will be renewed and some others will be only maintained with light maintenance. On the average on several years, the cost efficiency of the maintenance policy can be compared. To avoid unjustified subsidy variations, moving average can be used. Another possibility to skip this problem would be to admit that some infrastructure managers will bear cost of renewal periodically and that the equilibrium is to be reached over the entire life cycle cost.

The third remark has radical implications. Even if the possession time is introduced as an external factor which tends to introduce external heterogeneity in the comparison of maintenance policy, it remains that, if we consider only maintenance costs, we are able to capture only part of the

infrastructure manager’s performance. A low cost maintenance due to an excessively large possession time would be cost effective but may be not welfare improving because of the opportunity cost of the capacity lost. We do not encounter this problem with the Japanese yardstick competition mechanism. As the companies are integrated, they tend to optimize internally the trade-off between large possession time and more capacity offered to the trains. If the line is not saturated, the opportunity cost of the possession time is zero. Nonetheless, for the saturated parts of the network, this leads us to the next step: the capacity management.

4.4. The capacity management for railways infrastructure

We are now addressing the ability to perform some tasks which lead the infrastructure manager, possibly associated to other public bodies, to offer capacity. This encompasses the following tasks:

- optimization of capacity to maximize the production of slots;
- enhancement of capacity, including elimination of bottlenecks if desirable;
- slots allocation;
- traffic control.

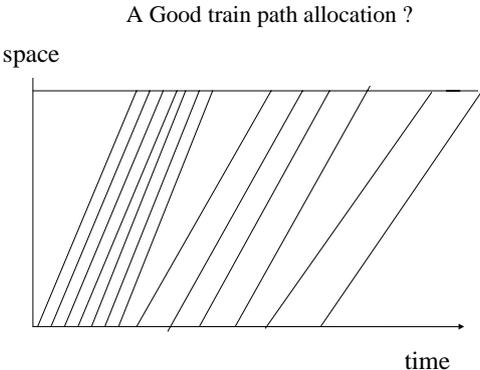
Implementing yardstick competition for the capacity management would be very difficult. Nonetheless, it is possible to make some comparisons that have a certain flavour of yardstick competition. We will see first the main problems.

First, we are facing different outputs: the capacity itself is an output as well as also the number of trains running on the network, but as the pricing system is out of the reach of the infrastructure manager, the resulting traffic is not the result of the effort of the infrastructure manager. Another output is the safety.

Another big difficulty is the heterogeneity of the train path. A freight train is generally slower and a passenger through train is naturally less capacity consuming than a slow train.

Last, but not least, the possibility to group together the slow trains and the fast trains can increase the capacity. In other words, the actual capacity, measured by the number of trains which can run on the line per day is not an objective per se, as the timetable has to be suited to the demand.(see figure 4.5)

Figure 4.5. Grouping of similar trains increases the capacity, but the slot allocation has to take into account the demand



We might add that the intervals between the trains depend partly on the infrastructure but they depend also on the technology of each train. So, even if the actual number of trains running on a line is to a certain extent the result of the effort of the infrastructure manager, it is greatly exogenous.

Theoretically, it should be possible to implement yardstick competition on a set of very similar lines, with similar trains. For example regional lines, with no interference of freight trains during the daily service. But the unit to be submitted to comparison is the line, and a lot of management costs are common to the different lines which are parts of the same network. So, the common costs, not clearly attributable to any line in particular, make uneasy any costs comparison between the lines.

If the lines are saturated, it is quite impossible to value the opportunity cost of the slots wasted because of mismanagement. So we propose to ignore this kind of costs, even if some comparison can lead to judge that an apparently saturated line can actually bear more traffic. This would be a qualitative assessment grounded on comparison, but as there is no clear incentive mechanism, we will not consider this comparison as yardstick competition.

If we consider now various networks, the accounts of the infrastructure manager can be described as follows :

subsidies = (capital costs) + maintenance costs + management costs - infrastructure charges

It has to be studied if management costs can be broken down into two components:

- a component linked to the length of tracks;
- a component linked to the traffic.

For the most part of them, those costs are labor force.

To implement yardstick competition, it is desirable to exclude capital costs and infrastructure charges.

For capacity management costs, some further studies are necessary, but very likely the unit cost has to be calculated by the ratio of the costs divided by the number of train kilometers and controlled for the length of the track, the composition of the traffic and broken down according to the type of traffic (i.e. regional lines not saturated, intercity lines, high speed lines, suburban lines,...). This yardstick competition for management costs has to be implemented for comparable infrastructure manager. By this we mean that the infrastructure manager to be compared should have to perform comparable tasks.

4.5. The production of transport services

First, it is necessary to point out that this paragraph does not concern road infrastructure. Moreover, for the infrastructure manager, the traffic is to be considered as given, and exogenous, that means that the traffic actually on the infrastructure is not under the only responsibility of the infrastructure provider. Indeed, we can wonder whether it is possible to implement a kind of yardstick competition for infrastructure provision of integrated railways companies. For sure, the production, as usually measured in terms of traffic units (ton-kilometer plus passenger-kilometer or any other linear combination of the two) is too much linked to external factors to provide a good measure of the effort of the companies. It is probably possible to compare the different traffic evolutions and to try to assess to what extent those evolutions are due to external or internal factors. Nevertheless any assessment of the performance of railways on the basis of the traffic could be strongly biased (Savignat and Nash, 1999). So it would be preferable to make the hypothesis that train kilometers and vehicle kilometers are the actual outputs of the railways companies.

For the integrated systems, it is possible to implement yardstick competition, as the Japanese system illustrates clearly. But can we implement yardstick competition on the maintenance costs alone with integrated companies and can we use vertically separated infrastructure companies with vertically integrated companies together in the same sample ?

If there is accountancy separation (as it is the rule in the EU), the maintenance department of the integrated company can be treated as an infrastructure company. There might be another system of

yardstick competition for the operating companies or the operating divisions of the integrated companies. If there is no accountancy separation between infrastructure maintenance and train operation, it seems difficult to implement yardstick competition. Some internal cost allocation rules can hinder a fair competition, thus the competition has to be between integrated companies. The example of Japan illustrates that such a competition is possible and is apparently improving efficiency.

Why should yardstick competition on infrastructure alone should be desirable ? More generally why should it be desirable to separate infrastructure costs from operation costs ?

- First, it is important that the maintenance costs are quite well known, to estimate access pricing, because infrastructure charges almost always incorporate marginal costs, maintenance costs being part of it.
- Second, it is important to compare different maintenance policies in order to assess their cost efficiency (for example integrated maintenance versus outsourcing).
- Third, the time scale of maintenance is different from the one of train operation; hence it is relevant to compare separately maintenance and operations.
- Finally, a good knowledge of optimal maintenance costs is necessary to know whether an infrastructure should be shut down or not.

To conclude on the production of traffic, we will offer two concluding remarks:

1. The production of traffic is not an output to be considered when measuring the efficiency of the firms, even if good traffic management can increase the traffic.
2. When it is possible, it is desirable to compare the efficiency of the maintenance policy of different firms, integrated or not.

Conclusions on yardstick competition in infrastructure provision

Infrastructure provision consists of the following activities:

- initial construction;
- maintenance;
- capacity management;
- production of traffic.

The first conclusion is that it appears that there are limited possibilities to implement yardstick competition for initial construction, notably because of the large heterogeneity of engineering works.

Yardstick competition can be implemented for maintenance works and could be useful for maintenance divisions of integrated companies.

Yardstick competition for capacity management can be implemented on similar networks to discover some capacity wasted but one more flexible kind of comparison can also produce some efficiency gains.

Finally, this analysis of infrastructure management leads us to the conclusion that franchising might not be very efficient because of life cycle cost. There are some risks of opportunistic behavior: to diminish the cost of maintenance during the franchise duration, to the detriment of more renewal costs, which could be born by another franchisee.

SUMMARY AND CONCLUSIONS

Regulation is necessary for infrastructure because of the main market failures described in chapter 1. By its very nature, infrastructure is a monopoly. Moreover, infrastructure management gives ample room to opportunism and might be qualified as subject to bounded rationality. Furthermore transport industry is producing numerous externalities.

Among the various options offered to the regulator, yardstick competition is particularly efficient. The main reasons for that are the following.

- First, it can be used in combination with other tools, among them notably:
 - franchising with the possibility to compare various franchised companies;
 - price cap with the possibility to set a price ceiling with reference to an average of costs, as in the Japanese system.
- Second, because this regulation mechanism is powerful in the sense that it provides strong incentives to reduce costs.
- Third, yardstick competition has the power to reduce significantly asymmetric information.
- Fourth, it seems also that the risks of collusion can be contained.

Moreover, yardstick competition has proved to be efficient under a certain number of circumstances in various industries: health care, transport, water supply and civil engineering. However, to the best of our knowledge it has not been used in the infrastructure management. But the Japanese yardstick competition mechanism includes an infrastructure cost comparison. It is also noteworthy to mention the fact that yardstick competition is, in Japanese railways industry, grounded on published and observable data. This might enable to reduce significantly the risks of capture.

From the previous analysis of infrastructure manager's activities, we can conclude that:

1. there is a limited possibility to implement yardstick competition for initial construction of transport infrastructure, the best way to introduce competition appearing to be competitive tendering;
2. maintenance can be regulated by yardstick competition, integrated firms can be included if their maintenance accounts are clearly separated;
3. regarding capacity management, even if some comparisons can help to find some productivity gains in slots allocations, it seems difficult to implement a pure yardstick competition, because of the strong heterogeneity of traffic running on the different lines. However, it appears that there is a possibility to implement yardstick competition on very similar lines.

Finally, we can conclude that yardstick competition might become one of the major tools of infrastructure provision regulation. It can be used as a part of a more complex framework, including price cap, cost plus or even ROR regulation.

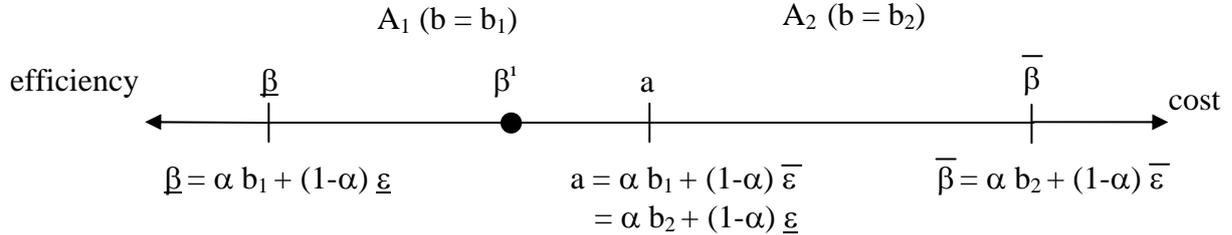
ANNEXE: AURIOL'S MODEL

Modelling productivity

We have already explained that the productivity parameter, β^i , consists of the two following elements:

- a common part, $b \in \{b_1, b_2\}$, which is the same for every similar firm of the market. This variable reflects statutory, regulation, organisational... particularities of the market. b_1 corresponds to high-efficient firms, whereas b_2 refers to low-efficient ones.
- a particular part, ε^i , which is assumed to be independent and identically distributed in a given interval. This individual characteristic includes the effects on the efficiency of firm i 's decisions (investments, management...).

Introducing the degree correlation between the firms, $\alpha \in [0,1]$, firm i 's characteristic is given by: $\beta^i = \alpha b + (1-\alpha) \varepsilon^i$ (the more the firms are correlated – α close to 1 –, the more the common part b is important within the characteristic). A high β^i corresponds to high cost, i.e., an inefficient firm. The authors usually assume that $\alpha b_1 + (1-\alpha) \bar{\varepsilon} = \alpha b_2 + (1-\alpha) \underline{\varepsilon} = a$, which ensures that both intervals $A_1 = [\underline{\beta}, a]$ and $A_2 = [a, \bar{\beta}]$ are disjoint⁷ (see figure below).



Theoretical results

Auriol (2000) proves that, in such a configuration, the use of comparisons improves the regulation efficiency. She considers N firms ; each of them has a cost function defined by $C^i = \beta^i - e^i$, where β^i is the productivity parameter described above, and e^i is the effort exerted by the firm. The regulator proposes the firms a revealing menu of contracts, defining the transfer T^i to firm i by the following:

$$T^i = a(\varepsilon^i) + \left[b(\varepsilon^i) C^i + (1 - b(\varepsilon^i)) \frac{\sum_{j \neq i} C^j}{N - 1} \right]$$

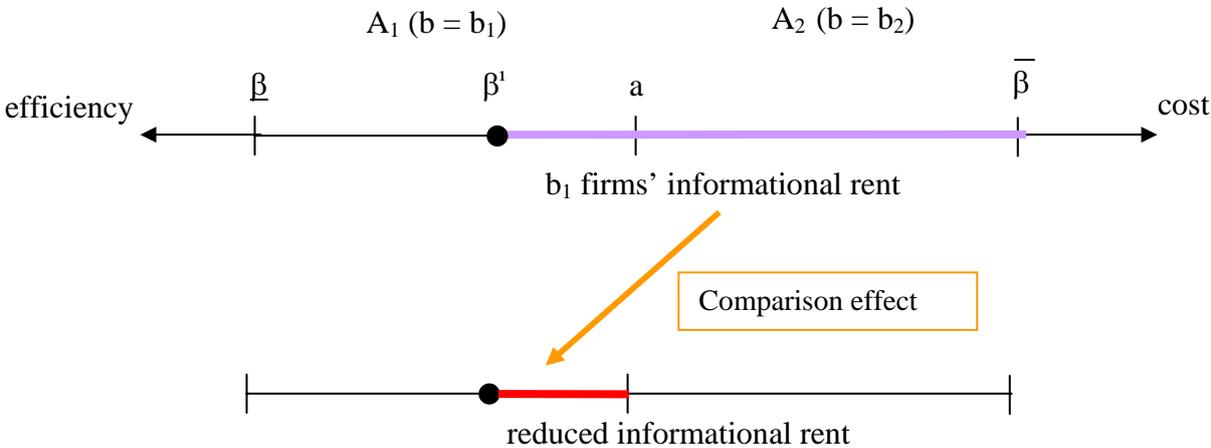
where $a(\varepsilon^i)$ is a fixed part, including the informational rent and the effort incentives ; $b(\varepsilon^i) \in [0,1]$ is the insurance against the risk that C^i exceeds the average cost. The functions $a(\varepsilon^i)$ and $b(\varepsilon^i)$ are so designed that, whatever $b \in \{b_1, b_2\}$ is, $a(\varepsilon^i)$ decreases with ε^i (the more the firms are efficient, the higher the

⁷ Auriol (1993) considers a duopoly structure with non-disjoint intervals.

informational rent and the effort incentives are), and $b(\varepsilon^i)$ increases with ε^i (the more the firms are efficient, the more they have to compete with their shadow firm). Using game theory, Auriol shows that firms are encouraged to choose the contract which corresponds to their productivity parameter, β^i . Hence, they reveal their common part, b_1 or b_2 . Knowing this information, the regulator is able both to reduce the informational rent of the high-efficiency firms (b_1), and to suppress the distortions on low-efficiency firms (b_2)'s production:

Reduction of informational rent

Following the above mechanism, the regulator can detect inconsistent announcements (when a firm announces a characteristic in A_2 , corresponding to a common part $b = b_2$, whereas the others announce characteristics in A_1). So, the regulator can retaliate against the firm of such an inconsistent announcement (excluding it from the market, for example), in order to make the mechanism revealing. So, a high-efficient firm cannot behave as if it were a low-efficient one anymore. The informational rent on the common part, b , is then suppressed. This truncates b_1 firms' informational asymmetry from $[\beta^i, \underline{\beta}]$ to $[\beta^i, a]$. The informational rent is reduced likewise:



Suppression of production distortions

Given that b_1 firms cannot pretend anymore to be of type b_2 , it is not necessary for the regulator to maintain the distortions on b_2 firms' production. Thus, the comparisons lead the regulator to make those firms exert the optimal effort level.

BIBLIOGRAPHY

- Auriol, E. (1993), "Monopole ou duopole : l'effet de comparaison", *Annales d'Economie et Statistique*, 31, pp. 1-31.
- Auriol, E. (2000), "Concurrence par comparaison : un point de vue normatif", *Revue économique*, 51, pp. 621-634.
- Averch, H. and L. Johnson (1962), "Behavior of the Firm under Regulatory Constraint", *American Economic Review*, 52-5, pp. 1053-1069.
- Baumol, W., J. Panzar and R. Willig (1982), *Contestable markets and the theory of Industry Structure*, Harcourt Brace Janovich, New York.
- Becker, G.S. (1983), "A Theory of Competition Among Pressure Groups for Political Influence", *Quarterly Journal of Economics*, 98, pp. 371-400.
- Bivand, R. and S. Szymanski (1997), "Spatial dependence through local yardstick competition : theory and testing", *Economics Letters*, 55, pp. 257-265.
- Bouf, D. and P.Y. Péguay (2001), "Is yardstick competition desirable for western European railways ?", *International Journal of Transport Economics*, 28-2, pp. 205-227.
- Carlton, D. and M. Perloff (2000), *Modern Industrial Organization*, Addison Wesley Longman.
- Coase (1945), "Price and Output Policy of State Enterprise", *Economic Journal*, 55, pp. 112-113.
- Dalen, D.M. (1998), "Yardstick competition and investments incentives", *Journal of Economics and Management Strategy*, 7-1, pp. 105-126.
- Dalen, D.M. and A. Gomez-Lobo (2003), "Yardstick on the road: regulatory contracts and cost efficiency in the Norwegian bus industry", *Transportation*, 30, pp. 371-386.
- Demsetz, H. (1968), "Why regulate utilities ?", *Journal of Law and Economics*, 11, pp. 55-65.
- DREES (2002), "La tarification à la pathologie, les leçons de l'expérience étrangère", *Dossiers solidarité et santé*, Hors série Juillet 2002, La Documentation Française.
- Hicks, J.R. (1935), "Annual survey of economic theory: the theory of monopoly", *Econometrica*, 3, pp. 1-20.
- Holmström, B. (1982), "Moral hazard in teams", *The Bell Journal of Economics*, 13, pp. 324-340.
- Improverail (2003), Deliverable 4, research project for the European Commission.
- Klein, B., R. Crawford and A. Alchian (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process", *Journal of Law and Economics*, 21, pp. 297-326.

- Laffont, J.J. and J. Tirole (1993), *A theory of incentives in procurement and regulation*, MIT Press.
- Mizutani, F. (1997), “Empirical Analysis of yardstick competition in the Japanese railway industry”, *International Journal of Transport Economics*, 24-3, 367-392.
- Nash, C. (2000), “Modelling performance: Rail”, in Hensher D.A. & Button K.J. (Eds.), *Handbook of transport modelling*, Pergamon, pp. 565-575.
- OFWAT (1997), *1996-97 report on water and sewerage service operating costs and efficiency*.
- Okabe M. (2004), “New passenger railway fares”, *Japan Railway & Transport Review*, 37, pp. 4-15.
- Preston, J. and C. Nash (1993), (“Franchising rail services” in Anthony Harrison (Ed.), *From hierarchy to contract*, Policy journals, pp. 147-165.
- Riordan, M. and D. Sappington (1989), “Second sourcing”, *Rand Journal of Economics*, 20-1, pp. 41-58.
- Savignat M-G., Nash C(1999), « The Case for Rail Reform in Europe--Evidence from Studies of Production Characteristics of the Rail Industry », *International journal of transport Economics*, 26(2): 201-17
- Shleifer, A. (1985), “A theory of yardstick competition”, *Rand Journal of Economics*, 20, pp. 41-58.
- Stigler, G. (1971), “Theory of Economic Regulation”, *Bell Journal of Economics and Management Science*, 2, pp. 3-21.
- Tirole, J. (1993), *The theory of industrial organization*, MIT Press.
- Van de Velde D., (2003), “Regulation and competition in the European land transport industry : some recent evolutions”, presented at the Thredbo 8 conference, Rio de Janeiro
- Vickers and Yarrow (1988), *Privatisation : an economic analysis*, MIT Press.
- White (2000), “Experience in the UK bus and coach industry”, in Bradshaw B. and Lawton Smith H. (Eds), *Privatization and Deregulation of Transport*, McMillan press.