ELSEVIER

International Journal of Transport Management 2 (2004) 5–15

INTERNATIONAL JOURNAL OF
TRANSPORT
MANAGEMENT

www.elsevier.com/locate/traman

European railway infrastructure: towards a convergence of infrastructure charging?

Yves Crozet *

Laboratoire d'économie des Transports (LET), University Lyon2, LET–ISH, 14 Avenue Berthelot, F 69363 Lyon Cedex 07, France

Received 1 January 2004; received in revised form 25 May 2004; accepted 25 May 2004

Abstract

Thirteen years ago, the European Directive 91/440 was implemented with the aim of increasing the efficiency of railway organisations throughout Europe. The main requirement of the Directive was to split transport operating from infrastructure management (at least in accounting terms). It also started the process of opening access for new operators; this is steadily being extended. Consequently, the issue of infrastructure charges is becoming more and more important. The paper underlines the fact that despite national differences, there are some signs of ways of convergence of infrastructure charges, due to the fact that the same key issue remains about infrastructure management: how to finance the renewal and development of infrastructure? So, among a wide range of objectives, some of them are becoming more and more crucial leading the principles of infrastructure charging to converge.

1. Introduction

Thirteen years ago, the European Directive 91/440 was implemented to increase efficiency of railway organisations throughout Europe. After many decades of continuous decline of rail modal share of market, this trend had to be stopped, because railway transport remains a desirable transport mode, especially due to environmental advantages compared to road transport. The main requirement of the Guideline was to split transport operating from infrastructure management (at least in accounting terms) and aimed at opening railway operation to competition and to develop market relationships in order to increase efficiency. Ten years after, the critical point of on track competition was still at the heart of the new European Directives (2001/12-13-14), first for freight traffic. Consequently, the issue of infrastructure charges is more and more important. According to the objectives of efficiency and competition, is there a best way to implement rail infrastructure charges? Could we find in Europe some "best practices" easy to transfer from one country to another?

In a first analysis, both of these questions have to be answered negatively. In the railway sector, infrastructure charges vary widely from one country to the other, in level as well as in the principles used for their calculation. While the same European required by have led to very different railway reforms across countries, the railroad infrastructure charging policies similarly differ a lot from one country to the other in as much as differences are numerous and noticeable in several fields. Following the first steps of the railroad reform between 1994 and 1997, the following variations could be mentioned.

- Some countries (Germany, Great Britain) had set up relatively high infrastructure charges, while others (France, Sweden, Switzerland) had deliberately set them at a low or even nil (The Netherlands) level.
- Some countries had opted for a high level of investment in infrastructures (Germany, France, Switzerland), while others were not doing any particular effort in that respect (Great Britain, The Netherlands).

By combining these two criteria, paradoxical situations appeared. While a country like Germany was following an economically logical path (high infrastructure charges and high investment level), others were on the

* Corresponding author. Tel.: +33-4-72-72-64-03; fax: +33-4-72-72-64-48.

E-mail address: yves.crozet@let.ish-lyon.cnrs.fr (Y. Crozet).

1471-4051/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.ijtm.2004.05.003

opposite choosing an apparently non-economic logic: high investment level combined with low charges (France, Sweden, Switzerland) or on the contrary high charges and low investment level (Great Britain).

Over time, such differences have been partially reduced. The level of railroad tolls have for instance increased a lot in France, for lines with high traffic. However, some divergences remain because many uncertainties still linger as regards best practices in this field.

- We will thus, in the first part present the theoretic foundations of these uncertainties. Indeed, if countries have troubles choosing the optimal charging policy, it is because the economic theory itself hesitates between several solutions. As required by the last European Directive, infrastructure charges should be based on the marginal social cost. But it is also possible to take into account long term considerations, especially the cost of investments.
- The second part will describe how some national railroad infrastructure charging policies bring their own answer to the infrastructure charging issue. These particularities will be acknowledged in a brief presentation of the evolution of British, German and French charging policies. This will lead us to stress the fact that there are a lot of goals within a charging system.
- From these national differences, the conclusion will underline the ways of convergence and especially the fact that some key issues remain about infrastructure management: how to improve the efficiency and how to finance the renewal and development of infrastructures? That is to say that among a wide range of objectives, some of them are becoming more and more crucial leading the principles of infrastructures charging to converge.

2. Theoretical recommendations about infrastructures charges: more issues than answers

It is generally agreed that the short run marginal cost (hereafter SRMC) constitutes the best theoretical solution to the issue of infrastructure charging. The principle is rather simple. SRMC pricing enables to run all trains for which the additional costs borne by society are inferior to the utility of that train for society, supposed to be represented by the willingness to pay. But even if the principle is simple, the question of what kind of costs have to enter in the SRMC is rather tricky. That is to say must we rather adopt a long run marginal cost (LRMC) principle? Another question arises about the best kind of tariff: single part or multi-part? And in that case, what could be the bases of varying the charge? Must we take into account investment costs necessary to develop the network? As we will see, there is no single and simple answer.

2.1. SRMC versus LRMC: from congestion cost to invesment cost

SRMC represents how the infrastructure costs change in the short-run when rail traffic levels change. The requirement for efficient pricing is to ensure that the costs borne by the user of the infrastructure reflect the sum of the marginal costs of the infrastructure provider, the infrastructure user and the others outside the transport system concerned.

- The marginal costs to the infrastructure provider in the short run are mainly the costs of wear and tear (as well as any additional costs of traffic control or signalling).
- Costs experienced outside the transport system are predominantly external accident and environmental costs. The price relevant marginal cost is thus the sum of these marginal costs less the costs in any case borne by the individual user.
- The variations in infrastructure cost with traffic levels are the short-run variable charges. They comprise several elements like track usage charges, administrative costs in running additional trains, peak charges and traction current charges, for instance electricity for traction purposes.

Even if, apparently, it is easy to add all these kind of costs, the more you increase the quantity of variables the more you complicate the calculation. And it is difficult to know precisely the relationship between the level of traffic and the variation of each part of the cost. Especially when you decide to take into account environmental costs and congestion costs.

- If environmental costs are included in SRMC for railways and, for example, not or not adequately for road trucking, this might lead to an undesirable competitive distortion, precisely causing what the decision maker wants to avoid: a pollution increase. Taking into account the accident costs exhibits the same conflicting alternative. Because of that difficulty, in many countries, external costs are not accounted for directly in the infrastructure charges. External costs of rail are globally compared to the external costs of road, but there is no specific part of the tariff for that. Such a solution does not conform to the economic theory, even if we adopt a ''second best'' point of view (see box 1). But according to the difficulties of an ex ante determination of external cost, the ex post calculation could be seen as a ''third best''. That kind of ''solution'' is not so obvious for congestion costs.

Box 1: Second best theorem

Efficient pricing of transport infrastructure is a necessary condition for maximising the social surplus. If the objective is to improve the use of the existing capacity, the short-run efficiency principle is based on given capacity. In the long run the efficiency condition is that investments in transport infrastructure should be undertaken up to the point where benefits just exceed costs (Small, 1992). However, this result supposes, inter alia, that the remainder of the economy is at the optimum, i.e. prices, at least in the sectors of the economy related to the sector studied, are equal to the marginal costs, which is often not the case.

Faced with a non-optimal situation in the other sectors, a marginal cost pricing only in one sector does not necessarily lead to an optimum in this sector. This established fact corresponds to the "second-best" theorem (Feldman, 1987). Precisely, for Lipsey and Lancaster (1956): "Given that one of the Paretian optimum conditions cannot be fulfilled, then an optimum situation can be achieved only by departing from all the other Paretian conditions".

This theorem of the "second best" thus seems to singularly weaken the theoretical prescription of marginal cost pricing. However that does not call into question the principle of pricing in itself. There is thus a consensus to judge that it is more efficient to charge something in particular for congestion and environmental externalities, rather than to charge nothing or to charge a price disconnected from the marginal cost (Raux and Souche, 2000). Concretely that means pricing that varies with the degree of congestion or nuisances (accidents, noise, pollution) emitted by the various transport modes.

- Per se, the congestion costs which can be decomposed into delays and opportunity cost of not getting a path are difficult to evaluate: what is the value of a delay for a particular train and what is the value of not running for a train? Moreover, the congestion costs might be borne, at least partly by the trains of the operator causing the congestion. In that case the operator is bearing the congestion cost twice. The theoretical solution is to give back the amount of the charges which is due to delays or opportunity cost of trains belonging to the same operator. But this might be the source of cumbersome calculation, sometimes rather arbitrary. Due to the difficulty of calculating the congestion cost, it seems easier to take into account the LRMC.

The concept of SRMC is often contrasted with LRMC, which represents the additional cost of an extra train when the infrastructure is optimally adapted to the demand in question. The general perception that short run marginal cost is below long run is only true in the presence of excess capacity; the reverse is true when capacity is scarce. In that case, instead of implementing a very high congestion tariff, it seems more convenient, according to the low capacity of payment, to adopt the LRMC principle. That is to say a definition of the tariffs grid including investment costs, and an administrative planning of the optimal number of trains on each part of the network.

LRMC is defined as the cost of an additional train 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. Long-run marginal costs represent the additional costs of providing an additional unit of traffic, under the condition that the level of infrastructure can be adjusted. We can distinguish long-run incremental costs, where the LRICs are the additional capacity and other costs of handling an extra block of traffic. These LRICs can be expressed in unit terms as long-run average incremental costs which are equal to the LRICs averaged over the extra units of traffic handled by the additional capacity. Precisely, this element of access charge represents the long run cost of maintaining and renewing the railway infrastructure. Although, the density of train operations over a particular part of the network is a significant factor in ascertaining the relative importance of this element within the total track access charge (Cole and Holvad, 1999).

"In practice, indivisibilities and the time lags involved in adapting infrastructure to volume mean that differences between short and long run marginal cost are likely. This has resulted in a vigorous debate regarding the relative merits of short and long run marginal cost pricing" (Nash et al., 2001). 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. On the contrary this charging system leads to charges more stable over time and thus facilitates the establishment of contracts between operators and infrastructure managers. Long stable contracts may justify specific investments such as rolling stock. So, there are some arguments in favour of LRMC, even if it would deprive the public of valuable services (the services with prices between short run marginal cost and long run marginal cost).

The long run marginal cost pricing approach can be use when there is a difference between short and long run marginal costs. In this case of over time, the long run approach gives a more stable value, the charge of the level of congestion is internalised. Linked to this is the fact that most of the European operators have negotiated contracts for a number of years, so they can justify specific investments in rolling stock or fixed equipment such as terminals. For Nash et al. (2001), "one solution might be to charge long term contracts on the basis of long run marginal cost, but to sell paths on the spot market at short run marginal cost".

But, one of the major drawbacks of marginal cost pricing (short run or long run) stems from the fact that railways are experiencing economies of scope, densities or scale. For that reason, marginal cost pricing does not fully cover the costs. One solution to this is a contribution from the State. However, some see this as dangerous in terms of the incentives to efficiency; others fear that might prove inadequate to fund an appropriate level of investment. Moreover, there is no non-arbitrary way to allocate the full cost of railways activity between services. Thus, to cover the non-allocated parts 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 to the desired target level. One of these methods is to set up a multi-part tariff which gives the opportunity to have at once a low level of variable cost of infrastructure and an efficient global amount of infrastructure charges.

2.2. From single-part to multi-part tariff

In a single-part pricing system based on marginal costs, the condition that total cost must be covered can be obtained with an additional charge, incorporated into the single price. This additional charge may be linear, digressive or progressive for the user or be based on the operator's willingness to pay (e.g. Ramsey rule). The charges can be differentiated according to the type of traffic; suburban passenger, mainline passenger, freight traffic. Furthermore, a difference can be made according to the time slot, the route and punctuality requirements.

A two or multi-part pricing system might separate the additional charges and the marginal costs. These additional charges can take into account:

- Speeds (defined according to optimal capacity utilisation);
- capacity situations on heavily-trafficked segments and at certain times (peak load pricing);
- the objective to cover the deficit arising with incremental cost pricing. This additional charge can be a fixed time-dependent contribution or might be linked to demand characteristics (e.g. high-speed traffic, regional traffic, combined traffic, etc.). Differentiation according to spatial parts of the networks is also possible.

The multi-part charging consists of a tariff with one fixed part and one or several variable parts, linked to the quantity consumed (by example train mileage). The interest of this formula is twofold.

It is possible to cover two objectives (at least), with on the one hand a variable cost near the marginal cost (goal 1), which ensure that the trains will be allowed to use the infrastructure to a degree commensurate to their social utility and on the other hand a fixed cost which is fixed according to a cost recovery target (goal 2). This makes it possible to run a train, whose "willingness to pay" (measuring the social utility) is only slightly bigger than marginal cost, the fixed cost being paid implicitly by other trains, for which the train paths are more valued.

It is also possible to propose a set of multi-part tariffs, with different fixed and variable costs. Facing this kind of "menu" the big operators will choose the tariff with big fixed costs and low variable cost. To the contrary, the small operators will tend to choose the tariff with low fixed cost and bigger variable cost. By a well design "menu" of tariffs it is possible to make the users of the infrastructure pay a bigger part of the surplus than with a single charging system (Baritaud, 2001).

Multi-part tariffs leave ample room for negotiation. Thus, as was observed for the Ramsey rule, there is a possibility of market power abuse, particularly towards captive customers (Baumol, 1983). To prevent or at least limit this risk, a rule is used in the US: the Stand-Alone Constraint (SAC). This criterion is devoted to limit, with a ceiling charge, the market power of a monopolist offering infrastructure usage to any user or group of users (Kessides and Willig, 1995). The ceiling is defined as the cost (indivisibilities included) that the user would bear if they were be the sole user of the infrastructure. This might be considered as the price that a hypothetical alternate infrastructure provider could offer. Thus, it represents a surrogate for a hypothetical competition.

The aim of this ceiling is clearly to limit abuse of market power. It is probably too soon to assess the potential utility of this kind of limitation in Europe. However, it is already possible to observe that in some very constrained situations, the market power is precisely linked to big indivisibilities. One example of that might be the main characteristic of European railway network which is not a unified one. Instead of one European network, we have the juxtaposition of national networks where one train operator was dominant. So, beyond the theoretical recommendations, we can also understand the divergences of infrastructure charges by studying the market power of the firms and the objectives of governments facing the dominant operator.

Box 2: Ramsey Principle

Alternative means of recovering more than simply the marginal cost of infrastructure use from rail operators, with the least possible damage to efficiency, have been proposed. The standard Ramsey pricing argument would justify raising price above marginal cost in inverse proportion to the elasticity of demand for the service in question.

In the case of public goods, this pricing principle allows collective welfare maximisation under a budgetary constraint. On the whole, 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 customer), the more the demand will be inelastic. So, in order to have higher tariff for operators 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 mark-up 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 criticised for its failure to protect captive customers. Moreover, as the tariff differs from the marginal cost, it is still possible that some services could be eliminated although their value is greater than their cost. It leaves some room for other kind of tariffs.

3. Infrastructure charging practices: which order for the goals?

Due to the great diversity of national practices in the field of rail infrastructure charges, it seems, at first sight, that logic behind them are very different. However, when looking at them more closely, it occurs that choices correspond to different ways of choosing the primary goals in stemming the charges.

We can easily point out six possible objectives for the infrastructure charging policy:

(1) favour the best possible use of the rail network;
(2) cover all (or part) of the operating and maintenance cost of the rail network;
(3) reflect the level of service provided to the carrier;
(4) contribute to the costs of developing the rail network;
(5) encourage the use of rail transport in intermodal competition;
(6) contribute to balanced regional development.

It is, however, clear that even a very sophisticated charging policy cannot aim at achieving these six goals simultaneously. On the contrary, everything leads to think that an order of priority will be set. Some goals will be forgotten or considered as secondary, while others will be favoured. It has to be noticed here that the economic analysis alone cannot solve this question. Different elements can explain the diversity in national choices. Geographical as well as political or social considerations play a non-negligible role in the set up of a charging system as well as in its evolution.

Indeed this is a second point we want to stress. Choices as far as charging policy is concerned are not straight forward. In most countries, there was a ''trial and error'' process going on. As a charging policy was relevant for one goal but less for one or several others, changes have taken place. The latter are probably not over yet.

We will thus try, for three countries, Great Britain, France and Germany, to present the reasons why such and such charging policy was chosen and then modified. In this respect, we will insist on the favourable as well as on the perverse effects of the chosen charging policy. To do so, we will stress the possible contradictions between the goals and how the three reviewed countries are facing deadlocks for one reason or another.

3.1. Great Britain: from privatisation and traffic increase to congestion and investment issues

During the reform of the British railroad system, another goal was added to the six previously presented ones and played a key role: John Major's *government will that the old historic operator British Rail be irreversibly broken up*. To do so, segmentation was total as far as management was concerned and privatisation was also pushed as far as possible, since Great Britain is the only European country that privatised its railroad infrastructure. The latter was then managed by the private company Railtrack, which, needless is to say, had to abide by certain technical (security, offer of paths) and charging constraints.

As far as charging policy is concerned, the British regulator applied techniques already used in other fields where privatisation and deregulation were taking place (water, phone, ...).

- As the "price cap" principle states it, access charges to infrastructure could not increase faster than a certain level defined by the classic formula "RPI + X". [1]
- As economic theory recommends it, infrastructure charges were divided in two parts, a fixed one for fixed costs and a variable one for variable costs. But the variable part was low in comparison with the fixed part. It has to be reminded that most of the Train Operating Companies (TOCs) received subsidies from the government and they were supposed to pay for infrastructure charges.

Results from this triple choice (infrastructure privatisation, price cap, bi-part tariff) were at first meeting the expected results. Nevertheless, they also, rather soon, faced certain limits that called into question the overall system's economy, which was obliged to change after a profound crisis.

- The first positive result was the substantial increase in traffic. Indeed, increasing the number of train per kilometre was the only way for the TOCs to achieve profitability. That was the case all the more as the charging system gave a relatively important place to fixed charges, independent of traffic. Average track usage cost was thus decreasing for the TOCs.
- The second result was growing traffic congestion on rails because of this train multiplication. As a result service worsens, which led Railtrack to propose a change in infrastructure charges in order to penalise operators that ran at busy times, thus causing congestion costs for the other users (especially in London's neighbourhood).

So, during the first stage of railway reform, compared to the six goals previously presented, Great Britain gave priority to the first (traffic increase) and second (covering costs) ones. Although it managed to achieve the first goal, it failed for the second one and experiences more and more difficulties with congestion and financing the renewal of infrastructures, resulting in the bankruptcy of Railtrack and its replacement by a 'not for profit' company, Network Rail. As a result, there has been a second stage in pricing policy, according to some recommendations of the Regulator (Office of the Rail Regulator, 2000; 2001). The new objectives were the following (Nash et al., 2003):

- an increase in the variable part of the track charges to reflect the full wear and tear cost and 50% of the quantified congestion cost;
- a move to a published tariff for all operators, with franchised operators continuing to pay on a two part tariff, but freight and open access operators paying only the variable element of the tariff;
- an incentive payment to Railtrack based on increases in traffic in order to encourage expansion of the network. In the same time, the new government decided an important amount of subsidies in favour of rail infrastructures.

All these modifications brought the structure of rail track charges in Britain much closer to the theoretical ideal. Wear and tear was more appropriately charged for, with the charges varying in fine detail according to the characteristics of the vehicle. A capacity charge was introduced based on the congestion cost caused by an additional train. But only 50% of the congestion cost was taken into account to avoid a too important reduction in service. With that last example, it is clear that in Great Britain, as in other countries, some conflicts remain between all the possible goals of pricing. The solution is probably, as in France, a development of congestion pricing on some specific part of the network only.

3.2. France: protection of historic operator and high speed trains financing

In France, the goals aimed at by the reform of the railroad system were totally different from those that prevailed in Great Britain. *While Great Britain explicitly tried to break apart the historic operator, France did everything to protect it*. That is why the separation between management (SNCF) and infrastructure (RFF) mostly enabled to transfer a large part of the former's debt to the latter. Concerning the involvement of RFF, the setting of a new pricing system of track access and track use was one the main task of RFF. But, the pricing system aims politically at removing the operational income deficit of the SNCF and achieving more or less balanced financial results.

At the beginning (1997), the charges paid by the SNCF were very low and based on a multi-part tariff in relation with a splitting of the national network in five areas: R(0) for peri-urban tracks (766 km), R(1) for high speed train tracks (831 km), R(2a) for high speed train track with moderate traffic, R(2b) for interuban track (4483 km) and R3 for other tracks (25,500 km). Three time-periods were then applied: peak period (6:30–9:00; 17–20:00), normal (4:30–6:30, 9:00–17:00, 20–0:30) and slack periods (0:30–4:30). Three tolling components were also taken into account:

[1] RPI + X = retail price index + a predefined element in the contract, that can be in some cases negative. In that case, price has to vary by less than inflation.

- DA for access rights which is a function of the distance in kilometres covered by the operating company.
- DC for operating which is a function of the train kilometres covered.
- DR for booking right, which depends on the time-period taken into account for R0 and R1 and on the train · kilometres covered.

After some slight changes, in 1999, the toll level for the French railway network was defined as indicated in Table 1. Key issues remained unsolved concerning the following questions:

- global level of pricing which was underestimated according to RFF;
- toll variation opportunities according to the level of service and quality of the network;
- the splitting of the pricing according to separate activities (and first between freight and passenger transport).

Nevertheless, RFF succeeded, step by step, to increase significantly the charges. Within that movement, the primary goal was not to weigh too much on SNCF's accounts so that it could present a balanced management account. Minor changes brought to this policy did not radically modify the system economy. Under the previously mentioned constraint, the latter led to differentiate a lot the infrastructure charges in space and time, depending on the service quality provided and on the manager's abilities to contribute. The result is obvious in Table 2. Without any important change in the structure of the tariff, the global amount of infrastructure charges raised, especially the charges paid by high speed train and urban trains. The first because of the capacity of payment of customers, the second because of the capacity of payment of local governments.

To understand the French charging system, two logics have to be clearly distinguished.

- The first one is explicitly an under-charging of infrastructure. For instance, the overall amount of charges paid by SNCF (single operator) to the network manager (RFF) was at the beginning (1997), half the maintenance infrastructure spending that the owner paid back to the SNCF, whose staff members are still in charge of the maintenance. Even with the 2003 tariff, infrastructure charges do not cover the total cost of maintenance. This could be explained by a willingness to foster railroad development or to contribute to regional development (goals 1, 5 and 6). This is however neither the only nor the principal explanation: abilities to contribute are also a major question.
- The charging structure, highly and more and more differentiated in time and space takes into account very closely the abilities to contribute as well as the service quality (goal 3). On most parts of the network, infrastructure charges are very low. On the contrary, they are much higher on two specific parts: urban zones and high speed lines. These charges should increase even more in the future. This discrimination is logical. Indeed, it aims at benefiting first from the capacity to contribute of the last-resort

Table 1
Tolling for track access and operating in France in 1999 (RFF)

	DA (€/km/month)	DR (€/km)			DC (€/km)
		Slack hour	Normal hour	Peak hour	
R0	1720	2.95	6.5	14.8	0.23
R1	1497	2.25	6	7.2	0.23
R2a	1497	2.25	4.5	6	0.23
R2b	6	0.33	0.33	0.33	0.23
R3	0	0	0	0	0.23

Source: RFF (1999).

Table 2
Tolling for track access and operating in France in 2003 (RFF)

	DA (€/km/month)	DR (€/km)			DC (€/km)
		Slack hour	Normal hour	Peak hour	
R0	373.12	1.55–0.62	4.97–1.24	14.38–2.49	0.806
R1	4475.91	4.813–0.806	9.78–1.68	11.54–3.2	0.806
R2a	3.11	0	0.08	0.08	0.806
R2b	0	0	0	0	0.806
R3	0	0	0	0	0.806

Source: RFF (2003).

payers: local governments for urban zones and firms for TGV profitability, which is assured by business customers. Then it enables to decide on the network's development. In places where investments are critical to increase capacity or to set up high speed trains, higher tolls are required. [2]

3.2.1. The French system thus mainly relies on a subsidy and under-charging logic

The French railway system (RFF + SNCF) receives about 10 billions of Euro subsidies per year, that is to say almost as much as its turnover. At that price, it was possible to develop tools that prove commercially and technically very successful (TGV) without questioning either the SNCF staff's situation or its monopolistic status. We can thus put forward that the main goal of the French system was to guarantee the existence of the historic operator. However, as this means very high subsidies, the charging policy of the infrastructure was built as a means to allow for developing the rail network (goal 4) in relationship with the requested quality of services (goal 3).

Considering the difficulties that both the British rail network was facing, the French choice seems relevant. However, it has to be kept in mind that it will rapidly end up facing two major difficulties:

- The first is of financial order since RFF faces increasing debts and manages to survive only because the French State regularly transfers large amount of money to the firm. In the meantime, the SNCF was again in deficit in 2001, 2002, 2003 and probably 2004.
- This problematic financial situation is not only due, as contended by the management, to increasing toll costs. The core of the problem lies in the meagre productivity gains of the firm. The latter will be a critical problem in the long-run. Indeed, since the SNCF is investing in other countries through its subsidiary firm (Keolis), in future years, it will have to accept in return to open up the French railroad to competition. As infrastructure tolls are quite low in France, compared to what prevails in Great Britain or Germany, competitors will quickly show better performances for the same service level.

Thus, the French charging system is also facing a true deadlock. It was created to protect the historic operator. However, as soon as competition will occur in the French railroad, for the freight first and then for passengers, this charging system will hurt the historic operator if it does not achieve noticeable productivity gains.

3.3. Germany: cost covering, competition and regional development

The German model of railroad transportation deregulation gathers two of the principal implicit goals of the British and French reforms. However, as it preceded both of them, these common points can be seen as constants in the deregulation process.

- From the conservative tradition, the German reform retained the idea that the infrastructure charging should cover operating costs, in order to avoid continuing subsidies. That is why the federal state became responsible for the railroad debt and the extra-cost linked to the status of the former Bundesbahn staff members. The latter, structurally transformed, was supposed to balance its accounts and privatisation was forecast afterwards.
- In the German logic of the national political economy (F. List), the idea was contended that deregulation should not threaten the existence of the historic operator even in its form of vertically integrated firm. As a result, the network manager (DN Netz) was maintained inside the holding company, DB (Deutsche Bahn).
- Contrary to what was done in France, competition was accepted in Germany and private operators were allowed to develop their activities, especially in the freight and regional passenger transportation sectors.

These different elements in hand, the chosen charging policy should not be too low. Any undercharging would have indeed led to a severe competition from new entrants, while it would have deteriorated DN Netz's hence DB's accounts. Germany could have then opted for a linear model, with a high charging differentiation as far as service quality was concerned. This was partially done in the first charging system, which proposed prices that varied not only with the train's technical characteristics (weight, length, speed, ...) but also with the quality of service required (precise or loose timing of the required slots ...). However, this logic was not fully developed because another strong constraint had to be accounted for by Germany: the reunification and hence the absorption of ReichsBahn. This railroad company used to hire in former Eastern Germany (16 million inhabitants) more staff members than the Bundesbahn in the western part (61 million inhabitants). Moreover, the network was obsolete and infrastructures were in bad shape. The federal State took charge of the renovation and of the development of new infrastructures in the East part. However, opposite to the French case, it

[2] Higher charges do not necessarily mean that costs are covered. Local government, as it was the case for the future Eastern TGV, can decide to subsidise the infrastructure, which reduces tolls by the same amount.

was not at stake to let users directly pay for the incurred cost, in the name of treatment equality of the different German regions.

According to the difficulty of determining a consistent single part and linear tariff, DB Netz, after a first revision in 1995 (reduction of the discounts which clearly favoured the DB companies), decided on a fundamental revision in 1998. A two-part tariff was established consisting of a fixed charge and a variable part per train kilometre. The aim was clearly to reduce the variable cost of use. But, in 2001, DN Netz decided another revision after a complain by competitors against the Antitrust Commission. They argued that a two part tariff is in favour of the main operator because of an implicit decreasing cost with the relative importance of traffic (Link, 2003).

Thus, the charging system of the use of infrastructure is again today a one part tariff, mainly linear system since the key element is the train kilometres run. This typically is a framework of regional development (goal 6) and of covering costs (goal 2). In view of the constraints, this choice seems rather logical. However, it was not able to solve all difficulties. Despite noticeable productivity gains, the DB is today in a very problematic financial situation and is facing increasing competition from private operators, which succeed better in regional transportation of people (e.g., Connex). In the same time, to meet the will of covering costs, tolls were set at quite high levels. This tended to prevent the development of railroad transportation as opposed to other modes of transportation like the road for the freight (goals 1 and 5).

The German railroad in thus facing a challenge, both financial and environmental, while the railroad system already received many large subsidies. From a certain standpoint Germany is ahead of Great Britain and France (noticeable productivity efforts) in its way of conducting reforms.

- Ahead of Great Britain because it has accepted for several years to heavily subsidise the renewing of the network.
- Ahead of France because noticeable productivity gains have occurred and are still realised nowadays.

Because of this relative advance, Germany is confronted with the very question of the 21st century railroad: is it necessary to develop railroad at all costs? Or shall it be circumscribed to its relevant area, which is today suburban transportation, high speed and freight, on some (dedicated?) lines only?

4. Conclusion: towards a convergence of tariffs objectives

The previous analysis of the three main countries would support some overall remarks. Although there is no clear cut line, it is possible to distinguish two broad tendencies: the short run marginal cost group (Great Britain, Switzerland, Norway, the Netherlands) and the long run one (France, Germany, Italy). Because of the countries' different objectives, the theoretical pricing principle can be used with local specificity. But it is clear that the six objectives cannot be reached at the same time. It is more and more obvious that whatever the rationale (SRMC or LRMC) three main questions still remain:

- First, the financing of infrastructures renewal and development that we can sum up by rising some questions: How to finance new infrastructures? Which way of subsidisation? And which effects on tariffs?
- Second, the risks of congestion on some parts of the network. Is it possible to reduce congestion by the way of wide modulation of charges in space and time? Is there the possibility to define a mix of ''Line by Line'' and ''Whole Network Cost Recovery'' framework?
- Third, the issue of efficiency. In the field of rail transport, as in the field of air transport, it is now clear that competition is not an objective per se. Competition is a means to reach higher efficiency. But because of increasing returns to scale, we are facing rather a competition between some giant firms than the pure and perfect competition between a great quantity of competitors. There are important barriers to entry within the railway system, and there are also sunk costs and other factors of monopolistic behaviour. So the question is: how to combine the necessary incentives of competition and the necessity of long term relationship between infrastructure manager and the main operator(s)?

It is not easy to find a scheme of infrastructure charges able to solve all these difficulties. It is nevertheless possible to indicate some main directions that many countries are now following. The convergence of rail infrastructure charges is in process: not in the sense of an equalisation of charges per train kilometre, but in the sense of an attempt to solve the same questions with some equivalent solutions.

- The first one is to maintain, and even sometimes to develop, public subsidies to rail infrastructures. But public subsidies shall be fair and relevant. The charging system shall provide the Infrastructure Manager (IM) and the operator with appropriate signals and incentives to invest. It is the reason why a wide modulation of tariff could be a good solution, especially when the subsidies are important. For the IM, this implies incentives to invest in new infrastructure. It is a well-known problem of SRMC pricing that it fails to give clear investment incentives. When applying a

SRMC pricing system the IM is likely to delay any investment in capacity development (e.g. an extra bridge or double track lines) because in a situation of a congested network the revenues can be higher than in the case of a network with little congestion. Clearly, from a macro-economic point of view, this will not be very wise as rail transport competes fiercely with road transport and timely delivery is one of the most important competitive edges.

- So it seems important to have a component of the tariff that enables the IM to raise the infrastructure charge as scarcity increases on a particular section. As long as the level of infrastructure charges does not exclude services from the market, or excludes only those services whose ability to pay would not be high enough to pay for capacity development, this scarcity-related price component brings additional revenues while keeping the incentive to carry out capacity development that are economically justified. Of course, regulatory controls are necessary to prevent the IM from taking advantage of its monopoly position by charging excessive prices while leaving congested areas unresolved. But more important in creating investment incentives is the establishment of long term contracts or investment engagements. That is to say that competition "off the track" is probably more important than competition "on the track". These mechanisms can directly solve capacity problems. By providing a balanced risk-sharing framework to invest, they ensure that all new infrastructure investment that can be paid back by profitable services is actually undertaken.
- It should also be kept in mind that State intervention gives important incentives for the IM to invest in new infrastructure. Especially, investment grant and interest reimbursement can give important incentives for an IM to build new infrastructure And it is why public subsidies could be, paradoxically, a good incentive to relevant investments. In fact in many countries a large proportion of the IM's income is directly provided by the State. [3] This means that it is the framework for determining these State reimbursements, rather than the infrastructure charging model itself, which determines whether the IM has incentives to achieve efficiency improvements. In those situations there seems to be little incentive originating from the charging system itself. In case the State reimbursement framework is the main efficiency driver, basically two incentives remain:
  - Reducing the budget: in terms of our model the portion of costs to be recovered by charging infrastructure fees will increase or the budget remains the same while the activities increase;
  - A too high infrastructure charge due to high levels of inefficiency will lead to a diminishing customer base and consequently the charge per customer will have to increase even further as the IM has to recover its costs. In contrast, low charges due to high efficiency levels will increase the IM's customer base.
- Finally, we would like to point out that apart from the actual charging system it is important that the actual contract between the IM and the operators also sets sufficient incentives for the IM to be cost efficient. Contract clauses whereby prices are revised based on achieved efficiency improvements will clearly diminish any incentives for improvement. On the other hand contracts whereby charges are set for a period of several years, possibly based on expected efficiency improvements over this period, can give a clear incentive. In this case, the agreed charges will not be affected by the actual level of cost reductions achieved by the IM for the length of the contract period. The IM will therefore have an incentive to reduce its costs and thereby to improve its return on investment.

[3] For example, this is the case in the Netherlands and Sweden, since there are no charges in the Netherlands yet and income from infrastructure charges in Sweden is passed directly to the State. Only in the UK the IM is fully responsible for generating a profit margin.

References

Baritaud, M., 2001. La tarification des infrastructures ferroviaires. Ph.D. Dissertation, Ecole des Mines de Paris.

Baumol, W., 1983. Some subtle issues in railroad regulation. International Journal of Transport Economics 10 (1–2), 341–356.

Cole, S., Holvad, T., 1999. Rail access pricing: an examination of the UK approach, in: Proceedings of the 8th WCTR, pp. 473–486.

Feldman, A.M., 1987. Welfare economics. In: Eatwell, J., Milgate, M., Newman, P. (Eds.), The New Palgrave. A Dictionary of Economics. Macmillan, London.

Kessides, N., Willig, R., 1995. Restructuring regulation of the railroad industry. World Bank Working Paper 1506, Washington, DC.

Link, H., 2003. Rail Infrastructure Charging and on-track competition in Germany—nine years later. In: Paper Presented at the Annual AET Conference, Strasbourg, October 2003, 16p.

Lipsey, R.G., Lancaster, K.J., 1956. The general theory of second best. Review of Economics Studies 24, 11–32.

Nash, C., Matthews, B., Whelan, G., 2001. Rail infrastructure charging in Europe—principles and practice. In: Proceedings of WCTR, Seoul.

Nash, C., Coulthard, S., Matthews, B., 2003. Rail track charges in Great Britain—the issue of charging for capacity. In: Paper Presented at the 8th International Conference on Competition and Ownership in Passenger Transport, Rio de Janeiro, September 2003.

Small, K., 1992. Urban transportation economics. Harwood Academic Publishers, New York. 181p.

Further reading

Bouf, D., 2002. Open access, vertical separation and infrastructure charging systems: railroad and railways. Global Business and Economics Review, Anthology, 310–321.

Bouf, D., Peguy, P.Y., 2001. Is yardstick competition desirable for western European railways? International Journal of Transport Economics 28 (2), 205–227.

Button, K., 1993. Transport economics. Edward Elgar. 269p.

European Commission, 1995. Toward fair and efficient pricing. Green Paper, 58p + annexes.

European Commission, 1998. Fair payment for infrastructure use: a phased approach to a common transport infrastructure charging framework in the EU. White Paper, 58p.

European Commission, 2001a. Guideline 2001/14 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification, Brussels.

European Commission, 2001b. European transport policy for 2010: time to decide (white paper). European Communities, 126p.

ECMT, 1998. User charges for railways infrastructure, ECMT.

Holder, S., 1999. Recent developments in rail infrastructure charging in the European Union. Journal of Transport Eco and Policy 33 (1), 111–118.

Jahanshahi, M.F., 1998. The US railroad industry and open access. Transport Policy 5, 73–81.

Nash, C., Matthews, B., 2002a. Implementing rail infrastructure charging reform-barriers and possible means of overcoming them. Institute for Transport Studies, University of Leeds, May 2002, Imprint—Europe Seminar, Brussels.

Nash, C., Matthews, B., 2002b. Implementing rail infrastructure charging reform-barriers and possible means of overcoming them. In: Paper for the Imprint—Europe Thematic Network, Leeds 2002.

NERA, 1998. An examination of rail infrastructure charges. Final Report for the European Commission, DG VII.

Nilsson, J.-E., 1992. Second-best problems in railway infrastructure pricing and investment. Journal of Transport Economics and Policy (September).

Nilsson, J.-E., 2000. Towards a welfare enhancing process to manage railways infrastructure access. Centre for Transport Economics, Working Paper.

Nilsson, J.-E., Fredriksson, J., Isacson, G., 1999. Organising a welfare maximising process to manage railway infrastructure access. In: Europe-Trip WP3 Market Game Model.

Quinet, E., 1990. Analyse économique des transports, Paris, PUF, 419p.

Raux, C., Souche, S., 2000. Acceptability factors to transport policy changes. In: European Transport Conference, Proceedings of Seminar C, Homerton College Cambridge, vol. P438, pp. 167–178.

Zoeteman, A., 2001. Life cycle cost analysis for managing rail infrastructure. European Journal of Transport and Infrastructure Research 1 (4), 391–413.