Optimal price and frequency on the railway passenger market

Etienne Billette de Villemeur
University of Toulouse (IDEI & GREMAQ).

September 2005
Preliminary - Comments are welcome - Please do not cite

Long Summary

Given the structure of the railway network, what is the optimal structure of services to be offered to rail travelers? More precisely, what is the (optimal) frequency of services and at what price should it be tariffed when one does take into account the complex network externalities that makes the flows inter-dependent? The paper is made of two parts in which we address theses question from three successive points of view.

In the first part, we restrict our attention to a single origin-destination pair in order to built a benchmark. We compute, for different objectives, the “optimal” price and frequency. in absence of any network externalities. The social optimum, where both travelers benefits and firms profits are taken into account, is considered first. Socially optimal prices equate marginal costs that can be considered null in the case of rail transportation. In other words, it would be optimal to give free access to train services (and compensate the companies for all their costs). Socially optimal frequencies are such that the travelers’ waiting time costs equate the average (production) costs.

We then consider the operator’s profit maximising strategies. Given the
particular cost structure (high fixed costs, no marginal costs), the profit-maximising strategies are no characterised in term of markup, but in term of price elasticity. More precisely, prices should be such that the price elasticity of demand equates one. The profit maximising frequency is again defined by a formula that links the average travelers’ waiting time costs with the average costs.

Lastly, we examine the second-best social optimum, i.e. the structure of services that maximizes social benefits and allows the operator to break-even. This is usually the objective that is assigned to public services.

The second part of the paper deals with the effect of network externalities on the optimal allocations introduced above. One of the specificity of rail is indeed that the characteristics of the different journeys are not independent from each other. More specifically, the price and frequency of services on one link directly impact the price and frequency of the journeys that encompass that link.

We consider a linear network and analyze the optimal strategies as a function of the demand for the various journeys. It appears that optimal prices do not necessarily refer to the (aggregate) price elasticities for the considered link, but to a weighted average of price elasticities for the various services that make use of that link. We identify in which circumstances it will be the case for the network that is considered and exhibit the optimal pricing formulae for the different objectives already introduced (social welfare, profit maximization and constrained social welfare).

Also the analysis is limited to the study a very simple network, it does highlight the specificities of the rail industry by taking into account the cost structure and the interdependency of the various flows. We thus believe that this is a first but important step toward a better understanding of the sector.

**Keywords:** Railways, passengers’ transportation, optimal pricing, optimal frequency, elasticity of demand, value of time, network externalities.
1 Introduction

“Revitalising the railways” is a posted objective of the EU. In particular, in 2000, the European rail stakeholders agreed “to increase its [the rail] market share of passenger traffic from 6 to 10%” by 2020. Over the last decades, however, the railways have not been able to face competition from other modes. Even if “volumes increased from 217 billion passengers×kilometers in 1970 to 290 in 1998”, “the market share fell from 10 to 6%” over the same period. By contrast, “The modal share of the car rose from 74% in 1970 to 79% in 1999.” And nowadays Air traffic volumes are similar to rail in terms of passengers×kilometers. There is no need to stress the importance of the environmental stake.

Environmental concern however is not the only reason for which the issue is of importance. Despite the slow decay of this transportation mode, the involvement of the public authorities in the rail system has always been important. “This state involvement however had, and still has, its price for the taxpayers in each EU Member State. State aids and other public contributions to the sector accounted for almost 40 billion € in 2001 in the EU”. Would it be for the sole public expenditures, it is worth to pay attention to the happening of the rail industry.

In order to take up the challenge of making train attractive, authorities point very much at efficiency improvement. There is however no clear view of what would be an efficient structure of rail services. More specifically, given the rail infrastructure, what should be the frequency of services and how should it be priced? This paper aims at answering, at least partially, these questions.

The optimal pricing of utilities is not a new issue. The first attempts to
address this question go back to the French engineer-economists school and in particular to J. Dupuit (1854). Various contributions explicitly deal with the rail industry (Dupuit, 1853; Hotelling 1938). Whatever the objective pursued (profits or social welfare maximisation), it appears that both marginal costs and price elasticities are essential to the computation of “optimal prices” (Lerner 1934, Boîteux 1956). If passenger rail transportation is considered, it should be noticed that the cost of conveying one additional passenger is almost negligible. This simple observation is the starting point of this paper. If indeed this is the case, the use of the standard pricing formulae becomes problematic. We will show that indeed, there is no simple explicit formulae for the optimal prices. However the later can be characterised by the value of demand elasticities. Interestingly enough, the predicted values that stream from our theoretical analysis are in line with most recent empirical studies (Oum et alii, 1990, Voith, 1991, 1997, Litman 2004) that attribute quite low values to the elasticity of demand with respect to fare, even in the long run.

The paper is divided in two parts. In the first part, we restrict our attention to a single origin-destination pair in order to built a benchmark. We compute, for different objectives, the “optimal” price and frequency. In absence of any network externalities.

The social optimum, where both travelers benefits and firms profits are taken into account, is considered first. Socially optimal prices equate marginal costs that can be considered null in the case of rail transportation. In other words, it would be optimal to give free access to train services (and compensate the companies for all their costs). Socially optimal frequencies are such that the travelers’ waiting time costs equate the average (production) costs.

We then consider the operator’s profit maximising strategy. Given the particular cost structure (high fixed costs, no marginal costs), the profit-maximising strategies are no characterised in term of markup, but in term of price elasticity. More precisely, prices should be such that the price elasticity of demand equates one. The profit maximising frequency is again defined by a formula
that links the average travelers’ waiting time costs with the average costs. Lastly, we examine the second-best social optimum, i.e. the structure of services that maximizes social benefits and allows the operator to break-even. This is usually the objective that is assigned to public services.

The second part of the paper deals with the effect of network externalities on the optimal allocations introduced above. One of the specificity of rail is indeed that the characteristics of the different journeys are not independent from each other. More specifically, the price and frequency of services on one link directly impact the price and frequency of the journeys that encompass that link.

We consider a linear network and analyze the optimal strategies as a function of the demand for the various journeys. It appears that optimal prices do not necessarily refer to the (aggregate) price elasticities for the considered link, but to a weighted average of price elasticities for the various services that make use of that link. We identify in which circumstances it will be the case for the network that is considered and exhibit the optimal pricing formulae for the different objectives already introduced (social welfare, profit maximization and constrained social welfare).

2 Price and Frequency on a Single Link

In this section, we construct a benchmark as to be able to study the impact of network externalities on price and frequency decisions in the next section. The section itself is split in three parts. The first part is devoted to respectively to the social optimum (or first-best) while the second exhibit the profit-maximizing strategies. The third part is devoted to the so-called second-best, i.e. the price and frequency that should be set by a benevolent (and perfectly informed) social planner that cannot make financial transfers to insure the survival of the rail operator, hence must insure that its decisions allow the later to break even.
2.1 The model

The transportation of $X$ passengers on the considered link gives rise to a gross surplus $S(X)$. Running $n$ train services induces for the firm a cost $nF$ that almost not depend on the number of passengers that are conveyed. We assume that, in addition, the operator must bear a fixed cost $F_0$. The fact that marginal cost may be considered to be null follows directly from train technical specifications. For the Eurostar High Speed train, with an empty weight of 752 tons, any additional passenger represents indeed about $1/10000$ of weight increase. For a city tram (electrical one-or-two coaches train), with an empty weight of about 50 tons, a passenger may represent up to $0.2\%$ of total weight. In all cases, marginal costs appears to be negligible as compared to average costs, hence to the price that an operator must charge to survive.

Denote $p$ the ticket price for the service under scrutiny. In taking their decisions, travellers balance the benefits they get from the service with all the costs they have to support. In addition to the financial costs, they also consider the time costs they have to incur. Denote $v_T$ and $v_W$ the value of time associated respectively with Travel and Waiting time. The travel time $T$ directly follows from technological constraints and is considered to be exogenously fixed by firm. The average waiting time between two services (that can also be interpreted as the average difference between the preferred departure time and the actual departure time) is inversely related to the number of connections $n$ that are offered on the link. Formally, we denote waiting time cost $v_W/n$ so that the generalised price (cost) to the consumers write:

$$\tilde{p} = p + v_T T + v_W \frac{1}{n}. \quad (1)$$

Given the notations introduced above, the aggregate demand $X$ is defined by:

$$X(p, n) = \arg \max_X \{S(X) - \tilde{p}X\}. \quad (2)$$

One can check that the demand function $X(p, n)$ have the expected proper-
ties \((\partial X/\partial p) < 0, (\partial X/\partial T) < 0\) and \((\partial X/\partial n) > 0\). Given the quasi-linear setup and the expression of generalised cost \(\tilde{p}\), it is possible to establish a relationship between those derivatives. More precisely, straightforward computations give (See Appendix 6.1.1)

\[
\frac{\partial X}{\partial T} = \nu_T \frac{\partial X}{\partial p},
\]

\[
\nu \frac{\partial X}{\partial n} = -\left(\frac{\nu W}{n}\right) \frac{\partial X}{\partial p}.
\]

### 2.2 Social optimum

The social optimum aims at maximising the difference between the benefits that result from services and the costs of all nature as supported by the firm and the passengers. At this point, the operator is not required to break-even. We thus implicitly assume that fixed costs can be financed at no efficiency cost through a subsidy financed from the general budget. Such a solution is usually considered not to be realistic and may even be explicitly forbidden. However, it constitutes an interesting benchmark.

Formally, social welfare writes:

\[
W(X, n) = S(X) - \left(\nu_T T + \nu_W \frac{1}{n}\right) X - nF - F_0.
\]

The optimal allocation is defined by the following two F.O.Cs:

\[
\frac{\partial W}{\partial X} = S'(X) - \left(\nu_T T + \nu_W \frac{1}{n}\right) = 0
\]

\[
\frac{\partial W}{\partial n} = \nu_W \frac{X}{n^2} - F = 0
\]

From (2), we know that the marginal utility of the representative consumer is equal to the generalised price of services so that equation (6) and (7) rewrite

\[
p = 0,
\]

\[
\left(\frac{\nu W}{n}\right) \frac{X}{n} = F.
\]
In words, ticket should be free and the number of services \( n \) should be such that the train cost \( F \) equals the waiting time costs of the \( X/n \) per train conveyed passengers.

Clearly, the allocation defined by (8) and (9) cannot be sustained since travellers should not bear any financial costs. In other words, the implementation of the social optimum requires government subsidies. It is usually not considered to be a realistic option. This is the reason why we focus later on the so called “second-best” allocation.

2.3 Profit-maximising monopolist

We now contrast the social optimum allocation to the situation that would prevail if the link under scrutiny is run by a profit-maximising monopolist. Given the setup introduced above, its profit write

\[
\Pi(p, n) = pX - nF - F_0. \tag{10}
\]

Differentiating with respect to both variables \( p \) and \( n \), one gets

\[
\varepsilon_p^X = 1, \tag{11}
\]

\[
p \frac{\partial X}{\partial n} = F. \tag{12}
\]

where \( \varepsilon_p^X \) is the (absolute value of the) price elasticity of demand as defined by

\[
\varepsilon_p^X = \frac{p}{X} \left( \frac{-\partial X}{\partial p} \right).
\]

Equation (11) is nothing but the standard Lerner formula when there are zero marginal costs. Interestingly enough, this “Lerner formula” does not say anything explicit about prices but makes a prediction on the value of the elasticity of demand. The latter however is a function of price. Equation (11)
thus says that the monopolist raise the price $p$ up to the point where the price elasticity of demand is equal to one.

By using equation (4) that links $(\partial X/\partial n)$ to $(\partial X/\partial p)$, equation (12) that defines the profit-maximising $n$ can be rewritten. Assuming that the price $p$ is indeed set to its profit-maximising level, (i.e. assuming that equation (11) holds), one obtains:

$$\nu W \frac{1}{n} = \frac{F}{X/n}. \quad (13)$$

Interestingly enough, equation (13) does not differ from equation (9) that defines the socially optimal level of $n$. As already pointed out by Billette de Villemeur (2004), the rule that governs the frequency of services is identical in both cases. Of course, this does not mean that the value will be identical in both cases. As expected, a profit-maximising monopolist will offer a lower number of services than what would be optimal.\(^5\) However, the distortion is somewhat limited in the sense that the difference in value proceeds solely from the difference in prices. There is no additional distortion from letting the firm setting the number of services.

### 2.4 “Second-best”

We now exhibit the strategies $(p, n)$ that maximize the social welfare $W$ and yet allow the firm to break-even. Let $L = W + \lambda \Pi$ denote the Lagrangian of the problem, where $W$ and $\Pi$ are respectively the welfare and the profit functions introduced above and $\lambda$ is the Lagrange multiplier.

Differentiating $L$ with respect to $p$ gives rise to the pricing rule

$$\varepsilon X_p = \frac{\lambda}{1+\lambda}, \quad (14)$$

which is nothing but the standard Ramsey formula when marginal costs are zero.

\(^5\)To see that, proceed by contradiction. Assume that equation (9) holds and that both value of $n$ are identical. Since $p$ is strictly positive when setted by a profit maximising monopolist, the demand $X$ is strictly lower than what would be socially optimal. Thus equation (13) cannot hold. In order for it to do so, it is necessary to decrease $n$. 

9
Not surprisingly, differentiating $L$ with respect to $n$, assuming that (14) holds and making use of (4) gives rise to the very same rule for $n$:

$$\nu_W \frac{1}{n} = F \frac{X/n}{n}.$$  

Again, it is clear that, in term of values, $n$ will be in-between the socially optimal value and the one that would be set by a profit maximising monopolist.

Interestingly enough, the pricing rule (14) may provide a first indication of whether there is to much or to little state involvement. If the operator makes about no profits but the (long run) price elasticity is lower than $\mu/(1 + \mu)$ where $\mu$ is the cost of public funds, then the marginal loss attached to the taxation system is higher than the welfare losses that follow from imposing the operator to cover all its costs trough ticketing. In other words, it would be more efficient to decrease the financial aid to the industry.

3 Network Externalities

We now consider the interaction between the different links that compound a network and its impact on optimum/equilibrium outcomes. As to identify the main mechanisms at work, we consider a very simple linear network. There are three cities $A$, $B$ and $C$. The demand for travel between $A$ and $B$ is denoted $X_{AB}$, the ticket price $p_{AB}$, the travel time $T_{AB}$, the number of trains $n_{AB}$ and their running cost $F_{AB}$. Notation is similar for the two other city pairs: $Q_{BC}$ and $Q_{AC}$.

The size of the cities a priori differ so that, even if distance between cities would be identical, the aggregate surplus $S_j$ cannot be assumed independent of the link considered $j \in \{AB, BC, AC\}$. By contrast, individual characteristics are assumed to be homogeneously distributed over the territory so that both value of time $\nu_T$ and $\nu_W$ are assumed to be representative for all type of passengers.
3.1 Omnibus services

We first consider that all trains run over the whole network (i.e. from A to C) and stop in B. As a result, \( n_{AB} = n_{BC} = n_{AC} \). In this simple context, the welfare now writes:

\[
W(X_j, n) = \sum_{j \in \{AB, BC, AC\}} \left[ S_j(X_j) - \left( \nu_T T_j + \nu_W \frac{1}{n} \right) X_j \right] - n(F_{AB} + F_{BC}) - F_0. \tag{16}
\]

The profits write

\[
\Pi(p_j, n) = \sum_{j \in \{AB, BC, AC\}} p_j X_j - n(F_{AB} + F_{BC}) - F_0, \tag{17}
\]

where the prices \( p_j \) must verify some obvious incentive constraints to which we shall come back later.

3.1.1 “First-best” over the network

Provided the optimal allocation indeed consists in offering “omnibus services”, the precise flows \( X_j, j \in \{AB, BC, AC\} \) and the number of trains \( n \) at the first-best optimum are defined by the following F.O.Cs:

\[
\frac{\partial W}{\partial X_j} = S'_j(X_j) - \left( \nu_T T_j + \nu_W \frac{1}{n} \right) = 0, \tag{18}
\]

\[
\frac{\partial W}{\partial n} = \nu_W \frac{\sum_j X_j}{n^2} - (F_{AB} + F_{BC}) = 0. \tag{19}
\]

Equation (18) just states that there should be no ticketing \( (p_j = 0, \text{ all } j) \). In other words, would it not create incentive problems, it would be optimal to subsidize the operator and finance all its costs. Equation (19) states again
that the average waiting-time costs should equate the average running cost. Formally

$$\nu_W \frac{1}{n} = \frac{F_{AB}}{\sum_j X_j/n} + \frac{F_{BC}}{X_j/n}. \quad (20)$$

Note that there is a potential conflict of interest between $AB$ and $BC$ travellers. There is indeed $a priori$ no reason to have equality of average costs ($F_{AB}/X_{AB}$) and ($F_{BC}/X_{BC}$). As a result, the very fact that the frequency of services is identical implies that, at the first-best, some travellers may get less services than what they would get if the services were run independently. By contrast $AC$ travellers always decrease the average running costs. Thus both $AB$ and $BC$ travellers have congruent interest with $AC$ travellers that constitute the driving force for the unity of the network. More precisely, at the first-best optimum, for the network to be beneficial for all passengers the two following participation constraints should hold:

$$\frac{F_{AB}}{X_{AB}/n_{AB}} \geq n \frac{F_{AB} + F_{BC}}{X_{AB} + X_{BC} + X_{AC}}, \quad (21)$$

$$\frac{F_{BC}}{X_{BC}/n_{BC}} \geq n \frac{F_{AB} + F_{BC}}{X_{AB} + X_{BC} + X_{AC}}. \quad (22)$$

The latter conditions rewrite

$$\frac{F_{AB}}{X_{AB}} \geq \frac{F_{BC}}{X_{BC} + X_{AC}}, \quad (23)$$

$$\frac{F_{BC}}{X_{BC}} \geq \frac{F_{AB}}{X_{AB} + X_{AC}}. \quad (24)$$

where the flows $X_j$ are considered for a zero price and an identical frequency.

3.1.2 Profit-maximization

Provided again the profit-maximising strategy indeed consists in offering “omnibus services”, the optimal strategies $(p_j, n)$ set by the monopolist are
defined by
\[
\frac{\partial \Pi}{\partial p_j} = X_j + p_j \frac{\partial X_j}{\partial p_j} = 0, \tag{25}
\]
\[
\frac{\partial \Pi}{\partial p_j} = \sum_{j \in \{AB, BC, AC\}} p_j \frac{\partial X_j}{\partial n} - (F_{AB} + F_{BC}) = 0. \tag{26}
\]

If equation (25) hold, that is if prices \( p_j \) are such that
\[
\varepsilon_{X_j}^p = \frac{p_j}{X_j} \left( \frac{-\partial X_j}{\partial p_j} \right) = 1, \tag{27}
\]
then (26) rewrites
\[
\nu_I \frac{1}{n} = \frac{F_{AB} + F_{BC}}{\left( \sum_j X_j \right) / n}. \tag{28}
\]

In words, as in all situations considered previously, the average waiting time costs should equal the average transportation costs. The congruence/divergence of interests evidenced in the “first-best” case are also still present. However, the “participation constraints” (23)-(24) have to be considered for different values of the flow since prices are non-zero and \textit{a priori} different. As pointed out in another paper (Billette de Villemeur 2005), these constraints are nothing but the equations that define the core of a network formation game. Although the game hence the mechanisms at hand are identical, the precise shape of these constraints depends on the value functions of the different players, \textit{i.e.} the objectives they are assigned.

Even if there is no risk for the network to burst \textit{i.e.} if the frequency of services is such that (23)-(24) holds true, one must verify that the prices \( p_j \) defined by (27) verify some obvious incentive constraints. In particular, one must verify that

- Passengers will not find it beneficial to split their ticket into several one, that is
\[
p_{AC} \leq p_{AB} + p_{BC}. \tag{29}
\]
• It is less expensive to buy a ticket for the journey than a more comprehensive ticket, that is

\begin{align*}
p_{AB} & \leq p_{AC}, \quad (30) \\
p_{BC} & \leq p_{AC}. \quad (31)
\end{align*}

These constraints lead to consider five different regimes, depending on the incentive constraints that are binding.

**In Regime A,** no IC constraint is binding and prices are defined by the implicit equation \( \varepsilon^n_{p X_j} = 1 \). The very fact of having a network does not modify the prices \( p_j \) that are exactly identical to the prices that would arise if the services were run in complete isolation. The sole impact of the network is the number of services \( n \) that \textit{a priori} differ from the amount of services \( n_j \) that would be adopted if the segments were run in isolation.

**In Regime B,** the IC constraint (29) is binding. In other words, would prices be set according to the rule (27), it would be less expensive for \( AC \) passengers to split their ticket and buy an \( AB \)-ticket and a \( BC \)-ticket. Thus the operator must tariff according to the rule \( p_{AC} = p_{AB} + p_{BC} \) so that it has only two price instruments. Such a situation may arise because the city \( C \) is a big center, while city \( B \) is of no interest so that the demand for \( AB \) services is negligible as compared to \( X_{AC} \) and \( X_{BC} \) so that \( p_{AB} \ll p_{AC} - p_{BC} \). In any case, if (29) binds, the FOC conditions of the profit-maximisation program actually faced by the operator write

\begin{align*}
\frac{\partial \Pi}{\partial p_{AB}} &= X_{AB} + p_{AB} \frac{\partial X_{AB}}{\partial p_{AB}} + X_{AC} + (p_{AB} + p_{BC}) \frac{\partial X_{AC}}{\partial p_{AC}} = 0, \\
\frac{\partial \Pi}{\partial p_{BC}} &= X_{BC} + p_{BC} \frac{\partial X_{BC}}{\partial p_{BC}} + X_{AC} + (p_{AB} + p_{BC}) \frac{\partial X_{AC}}{\partial p_{AC}} = 0.
\end{align*}
Thus the two prices are defined by the implicit formulae

\[ \varepsilon_{AB} = \frac{X_{AB}}{X_{AB} + X_{AC}} \varepsilon_{p}^{X_{AB}} + \frac{X_{AC}}{X_{AB} + X_{AC}} \varepsilon_{p}^{X_{AC}} = 1, \quad (32) \]

\[ \varepsilon_{BC} = \frac{X_{BC}}{X_{BC} + X_{AC}} \varepsilon_{p}^{X_{BC}} + \frac{X_{AC}}{X_{BC} + X_{AC}} \varepsilon_{p}^{X_{AC}} = 1. \quad (33) \]

This pricing rules merit a few comments. First, prices should be set by considering the \textit{average} elasticities over both segments (that is \( \varepsilon_{AB} \) and \( \varepsilon_{BC} \)) rather than by considering \textit{aggregate} elasticities over both segments. Indeed the later write

\[ \varepsilon_{AB} = \frac{p_{AB}}{X_{AB}} \left( \frac{-\partial X_{AB}}{\partial p_{AB}} \right) = \frac{X_{AB}}{X_{AB} + X_{AC}} \varepsilon_{p}^{X_{AB}} + \frac{p_{AB}}{p_{AB} + p_{BC}} \frac{X_{AC}}{X_{AB} + X_{AC}} \varepsilon_{p}^{X_{AC}} \quad (34) \]

\[ \varepsilon_{BC} = \frac{p_{BC}}{X_{BC}} \left( \frac{-\partial X_{BC}}{\partial p_{BC}} \right) = \frac{X_{BC}}{X_{BC} + X_{AC}} \varepsilon_{p}^{X_{BC}} + \frac{p_{BC}}{p_{AB} + p_{BC}} \frac{X_{AC}}{X_{BC} + X_{AC}} \varepsilon_{p}^{X_{BC}} \quad (35) \]

where \( X_{AB} = X_{AB} + X_{AC} \) and \( X_{BC} = X_{BC} + X_{AC} \) denotes the number of passengers over both segment of the tracks. It is clear that the \textit{aggregate} elasticities are smaller than the \textit{average} elasticities. Thus, if the first are considered, the services are over-priced. Conversely, if the operator is indeed profit-maximising, the observed prices should lead to \textit{aggregate} elasticity estimates that are smaller than one. Second, in contrast to what for what happens for the number of services, both \( AB \)- and \( BC \)-travellers would prefer \( AC \)-travellers not to be there for the prices that emerge. Indeed, if (29) is binding, this means that prices \( p_{AB} \) and \( p_{BC} \) as defined by (32)-(33) are higher than those that would prevail if they were defined according to (11), that is if both segment were isolated. Finally, note that the pricing formula (32)-(33) are those to be considered whenever \( p_{AC} = p_{AB} + p_{BC} \) holds true. In other words, if the company impose itself this rule (e.g. for say the legibility of the tariff principles or in order to reduce the complexity problem that would follow from computing optimal prices for all possible journeys over the network), prices should obey the very same equations, whether the IC constraint (29) is binding or not.
In Regime C, one of the IC constraints (30)-(31) is binding. This means that for some passengers it would be cheaper to buy a comprehensive ticket. This may occur if, say, there is an intense competition for AC-travellers between rail and air-transportation companies. If (30) is binding, then \( p_{AC} = p_{AB} \) which is defined by

\[
\frac{\partial \Pi}{\partial p_{AB}} = X_{AB} + p_{AB} \frac{\partial X_{AB}}{\partial p_{AB}} + X_{AC} + p_{AC} \frac{\partial X_{AC}}{\partial p_{AC}} = 0.
\]

In other words, while \( p_{BC} \) is defined by (27), that is \( \varepsilon_{X_{BC}}^p = 1 \), the price \( p_{AB} \) is now defined by (32), that is \( \varepsilon_{AB}^p = 1 \). Note however, that when \( p_{AB} = p_{AC} \), the aggregate elasticity over the segment \( AB \) does not differ from the average elasticity. Indeed

\[
\varepsilon_{AB} = \varepsilon_{AB} - \left( 1 - \frac{p_{AB}}{p_{AC}} \right) \frac{X_{AC}}{X_{AB} + X_{AC}} \varepsilon_{AC}^p = \varepsilon_{AB}.
\]

In this regime, AC-travellers would prefer the train not to stop at \( B \) in order to benefit from cheaper tickets while \( AB \)- or \( BC \)-travellers benefits from a decrease in the ticket price (as compared to what they would pay if services were run in isolation).

In Regime D, both IC constraints (30)-(31) are binding. In this case, there is only one price to be set, which does not depend on the journey: \( p_{AB} = p_{BC} = p_{AC} = p \). If the later equality holds true, then the price \( p \) is characterised by:

\[
\frac{\partial \Pi}{\partial p} = \sum_{j \in \{AB,BC,AC\}} X_j + p \frac{\partial X_j}{\partial p} = 0.
\]

In other words, the aggregate elasticity \( \varepsilon = \left( p/\bar{X} \right) \left( -\partial \bar{X}/\partial p \right) \) where \( \bar{X} = \sum_j X_j \), that does not differ from the average elasticity

\[
\bar{\varepsilon} = \sum_{j \in \{AB,BC,AC\}} \frac{X_j}{X_{AB} + X_{BC} + X_{AC}} \varepsilon_{X_j}^p
\]

should be set equal to one.
With a linear-pricing rule: We know that whenever one of the constraints (30)-(31) is binding, the IC constraint (29) cannot be binding. Thus, there are no other regimes to be considered. However, it is often the case that for reasons that are completely orthogonal to the incentive issues, the operator does adopt a linear pricing scheme\(^6\). Consider thus that \(p_j = kl_j\) where \(k\) is the per kilometer price and \(l_j\) stands for the length of the journeys (in km) over \(j \in \{AB, BC, AC\}\). By definition, \(l_{AB} + l_{BC} = l_{AC}\). Profit maximization writes

\[
\frac{\partial \Pi}{\partial k} = \sum_{j \in \{AB, BC, AC\}} l_j \left( X_j + p_j \frac{\partial X_j}{\partial p_j} \right) = 0,
\]

that is

\[
\sum_{j \in \{AB, BC, AC\}} \alpha_j \varepsilon_{X_j} = 1,
\]

where \(\alpha_j = l_jX_j/\left(\sum_n l_nX_n\right)\) is the relative weight of type \(j\) passengers in terms of revenue or passengers \(\times\) kilometers. Of interest will be the link with the aggregate/average elasticity:

\[
\hat{\varepsilon}_k = \frac{k}{X} \left( -\frac{\partial \tilde{X}}{\partial k} \right) = \frac{k}{\sum_n X_n} \sum_{j \in \{AB, BC, AC\}} \left( -\frac{\partial X_j}{\partial k} \right) = \frac{1}{\sum_n X_n} \sum_{j \in \{AB, BC, AC\}} \frac{X_j}{X_n} = \bar{\varepsilon}.
\]

Finally, note that, although we do not derive them explicitly, equation (26) take different forms depending on the regime or pricing system that is adopted.

---

\(^6\)As already mentioned above, a simple pricing scheme may be adopted because the operator does not have enough information to compute the optimal prices, or find it too complex. It may also be the case to improve the legibility of the pricing scheme toward consumer, hence the acceptance of the tariffs.
3.1.3 “Second-best”

Turning back to the maximization of the social welfare $W$ as defined in (16) under the budget constraint $\Pi \geq 0$ gives rise to the implicit pricing rule

$$\varepsilon_{Xj}^p = \frac{\lambda}{1 + \lambda},$$

(36)

which is identical to the pricing rule (14) that prevail if segments are considered in isolation. Not surprisingly, if (36) holds true, the optimal number of services is defined by

$$\nu_W \frac{1}{n} = \frac{F_{AB} + F_{BC}}{\left(\sum_j X_j\right) / n},$$

(37)

which is identical to (20) and (28). This shows that it is a common rule for the network considered as long as the “incentive constraints” (29)-(31) are not binding. Again, for the formation of the network to be beneficial to all travellers, the “participation constraints” (23)-(24) should hold true. And as already underlined for the profit-maximising allocation, although the formulation of the constraints does not depend on the objective, there exact value does.

The four regimes identified above find again a natural place in the analysis. However, it is fair to say that the only change between the profit-maximising prices and those that correspond to the second best allocation is the value of the price elasticity of reference. More precisely, under Regime A (No IC constraints is binding), the prices are defined by the rule (36) just identified. Under Regime B, (the constraint (29) is binding), the prices are defined by

$$\varepsilon_{AB} = \frac{\lambda}{1 + \lambda} \quad \text{and} \quad \varepsilon_{BC} = \frac{\lambda}{1 + \lambda}.$$  

Under Regime C, assuming that (30) is binding, (that is $p_{AC} = p_{AB}$), the socially optimal prices are defined by:

$$\varepsilon_{AB} = \frac{\lambda}{1 + \lambda} \quad \text{and} \quad \varepsilon_{X_{BC}}^p = \frac{\lambda}{1 + \lambda}.$$  

18
Under Regime $D$, the aggregate/average elasticity should verify

$$\bar{\varepsilon} = \varepsilon = \frac{\lambda}{1 + \lambda},$$

while with linear tariffs, the socially optimal prices are defined by:

$$\sum_{j \in \{AB, BC, AC\}} \alpha_j \varepsilon p_{X_j} = \frac{\lambda}{1 + \lambda}.$$

Beyond the distinction between the different pricing regimes, of interest is again the comparison between the relevant elasticities and the ratio $\mu/(1 + \mu)$ where $\mu$ is the costs of public funds.

4 Conclusion

By taking seriously into account the specificities of the Rail industry, this paper deliver some interesting insights on the optimal structure of passengers’ services in the rail industry. First, since there is about zero marginal cost of conveying an additional passenger, it would indeed be optimal to give free access to rail services if this would not create incentive problems. It that sense, the paper provide some justification for State involvement in the industry. Second, if the operator is profit-maximising and there is no regulation (e.g. because inter-modal competition is considered to be sufficient to reduce market imperfections), the optimal price can no more be characterised by the means of the standard Lerner formula. The later is degenerate and makes now a prediction on the value of the price elasticity of demand. Similarly, under a regulatory regime, the optimal prices are characterised by the value of the elasticity of demand. Third, by measuring these elasticities and comparing it to a simple function of the costs of public funds, one may get an indication of wether or not there is too much State involvement within the industry. Forth, the rule for setting the frequency does not depend on the precise goal which is pursued. In all cases, it should be such that the cost of waiting equate
the average (production) cost. This says that, a) since travel time is not con-
sidered, passengers have always to support a larger generalised price than
the average cost, and b) there is no specific distortion introduced by the fact
that the operator may choose the frequency of services. Five, unless there is
a sufficient amount of travellers that make use of the different parts of the
network, there will always be some segment that find beneficial to be run
in isolation. In other words, long-distance travellers constitute the driving
force for the unity of the network. Six, generically, pricing cannot be consid-
ered journey by journey. Seven, due to the incentive constraints that limits
the pricing possibilities and/or due to ad hoc simple (linear) pricing rules,
generically, optimal prices are not characterised by aggregate elasticities but
by an adequate weighted sum of elasticities.

Of course, the papers bears a number of limitations. Although the frame-
work is almost simplistic, we fail to explore all its possibilities. In particular,
we did not consider the possibility that some direct services are offered for
long-distance journeys. Moreover the structure of the network is taken as
given and no consideration is made of issues like congestion (that may limit
the possibility just mentioned) or investment. These problems are left for
further research. Despite these obvious limits, we nevertheless believe this
paper constitute an interesting first-step that may be useful in the search for
an efficient use of the Rail infrastructures.

5 References


Caps”, *Transportation Research Part E*, 40(6), 465-476.

consequences of (optimal) network allocations”, *mimeo*.


Finn Jørgensen and Pål Andreas Pedersen [2001]. “The Influence of Travel Distance and Transport Operators’ Objectives on Fares, Transport Quality and Generalised Transport Costs”, *Discussion paper 01/04*, University of Kent.


6 Appendix

6.1 Calculus and proofs for the first section

6.1.1 Proof of equations (3) and (4)

From (2), we know that

\[ S'(X) = \tilde{p}. \]

Differentiating with respect to \( p, T \) and \( n \) gives

\[
\begin{align*}
S''(X) \frac{\partial X}{\partial p} &= 1, \\
S''(X) \frac{\partial X}{\partial T} &= \nu_T, \\
S''(X) \frac{\partial X}{\partial n} &= -\nu_W \frac{1}{n^2}.
\end{align*}
\]

Equations (3) and (4) follow directly.