TI 2001-109/3
Tinbergen Institute Discussion Paper

Choice of Frequency and Vehicle Size in Rail Transport: Implications for Marginal External Cost

Piet Rietveld

Department of Spatial Economics, Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, and Tinbergen Institute

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Choice of frequency and vehicle size in rail transport:
Implications for marginal external costs

October 2001

Piet Rietveld
Jasper Dekkers
Stefan van Woudenberg

Faculty of Economics
Vrije Universiteit
Boelelaan 1105, 1081 HV
Amsterdam
The Netherlands
fax: 31-20-4446004
email: prietveld@econ.vu.nl

Abstract.

Frequency of services and vehicle size are important policy instruments of railway companies. Extending Mohring's basic 'square root model' for frequencies, we arrive at more general formulations for frequency, vehicle size and price under alternative regimes of welfare and profit optimisation. It appears that in the more refined models the frequency response of railway companies with respect to changes in passenger volumes is not far removed from the standard square root result. After the formulation of theoretical models we also carry out an empirical analysis where choice of frequency and vehicle size are analysed with special attention paid to the level of occupancy rates. It appears that average occupancy rates are low in rail transport. As a result the frequency and vehicle size response of the railway company appears to be low. It is estimated that an increase in the number of passengers of 1% leads to an increase in the supply of capacity of about 0.5 % (a frequency increase of about 0.35% and an increase of vehicle size of about 0.15%). This has important implications for the environmental costs of railway firms. Given the low occupancy rate, an additional passenger does not lead to a proportional increase in capacity so that the marginal costs are lower than the average costs.

1. INTRODUCTION.

Producers of public transport services face decision problems with a considerable number of dimensions, including network structures, pricing, spacing of lines and stops, frequency of service, and vehicle size. In the present paper we will focus on the latter two aspects: choice of frequency of service and of vehicle size.

Public transport services are usually characterised by economies of vehicle size: costs per seat tend to decrease with increasing size (Mohring, 1976). Thus, cost oriented public transport firms would be expected to employ large vehicles. However, from the perspective of the passengers, such a policy would be unattractive since it would lead to low frequencies and hence to high schedule delay costs and waiting times, and also to low speeds given the time losses for boarding and alighting. This leads to the question how public transport companies will weigh the various interests concerned. In this paper we will discuss extensions of Mohring's well known square root principle which says that frequency increases with the square root of demand.

An important feature of public transport is that outside the peak periods there is substantial excess capacity. Even during peak periods average occupancy rates are often below 50% (Oldfield and Bly, 1988). One might wonder how this can be explained, because one would expect a proper planning of transport services in terms of frequencies and vehicle size to lead to more efficient operations. In this paper we will discuss reasons why occupancy rates often are so low in public transport. We will also pay attention to the implications of low occupancy rates for marginal and average cost of transport. As will be discussed in the paper, the occurrence of low occupancy rates has important implications for the structure of external costs. We will demonstrate that since occupancy rates are low in public transport the average external costs of the services are high, but the marginal costs are considerably lower.

2. CHOICE OF SERVICE FREQUENCY AND SIZE OF VEHICLES IN PUBLIC TRANSPORT.

The base model of frequency choice by public transport operators has been formulated by Mohring (1972). It can be outlined as follows. Consider the demand for trips per time period on a certain line (denoted as Q) as given ($Q=Q_0$). Further, let F denote frequency of service per time period and let costs of service equal

$$C_{operator} = u.Q+v.F \qquad (1)$$

where u is the marginal cost per passenger and v is the cost of an extra vehicle used to serve passengers[1]. In addition to the costs experienced by the operator there are also costs for the passengers related to waiting time and schedule delay. When the vehicles are equally spaced, the interarrival time between vehicles equals 1/F. This implies that the average waiting time for a traveller going to a public transport stop without consulting the time table equals 0.5/F. Since waiting time at platforms is valued more negatively than in vehicle time, the value associated to the waiting time is relatively high (say a factor a_1, where a_1=1). Then, when the value of travel time is denoted as a_2, the factor $a=a_1.a_2/2$ translates the inter-arrival time into monetary terms:

$$C_{traveller} = a.Q/F \tag{2}$$

Minimising the sum of total costs of company and travellers

$$C = C_{operator} + C_{traveller} = \quad u.Q + v.F + a.Q/F \tag{3}$$

leads to the optimum frequency:

$$F^{*} = [a.Q/v]^{0.5} \tag{4}$$

This result is known as the *'square root principle'*. It means that an increase of demand Q with 10% leads to an increase of frequency of services of 5%. In a similar way optimal frequency will respond positively to changes in the cost of waiting time per passenger (factor a) and negatively to changes in costs of supply of an additional vehicle (factor v).

One of the limitations of this result is that vehicle size is not considered explicitly: it is assumed to be given. This has implications for the occupancy rate OR defined as:

$$OR = Q/[F.S] \tag{5}$$

where F.S is the total capacity. Since the occupancy rate can be rewritten as:

$$OR = [Q.v/a]^{0.5}/S \tag{6}$$

it appears that it will increase with increasing demand, the pertaining elasticity being 0.5. Hence the square root principle means that with high demand the occupancy rate will rise above 1[2].

Figure 1. The relationship between capacity and travel demand according to the square root principle.

When one is interested in occupancy rates, this is a somewhat implausible result, since one would expect that at higher volumes of travellers the operators would respond by increasing the size of the vehicles. This obviously calls for a joint analysis of choice of frequency and vehicle size by operators.

Another point that deserves attention is the possible response of travellers to higher frequencies. In the base line approach demand is inelastic ($Q=Q_0$), but in a more general setting one would expect that travellers respond to higher frequencies and that operators take this into account in their decision whether or not to increase frequency.

Jansson (1980) introduced the issue of vehicle size by formulating a model where operators jointly maximise size and frequency, and where peak and off-peak periods are distinguished. Based on the assumption of inelastic demand he derives optimal levels of frequency and size of buses. The assumption is that during the peak the occupancy rate is 100%, whereas it may be lower at other times. He concludes that at the time of research the structure of bus operations in Sweden was clearly sub-optimal since frequencies were too low and bus size was too large. The explanation of this gap between the actual and the optimum outcome is the neglect of user costs by public transport operators.

Using computer simulation techniques, Glaister (1986) analysed the potential consequences of deregulation of public transport in the city of Aberdeen based on the assumption of loss minimising operators, and where also bus fares are taken into account. His conclusions are comparable to those of Jansson that at that time busses were too small and that frequencies were too low. Although deregulated bus companies would not take into account directly the user costs of travellers, they may yet benefit from higher frequencies when travellers are prepared to pay higher fares. One of the issues he raises is the possible emergence of differentiated services for different types of travellers, a point that has been investigated in more detail by Gronau (2000) who analyses optimum diversity in terms of service frequencies and vehicle size.

Oldfield and Bly (1988) formulate a model with elastic demand where social benefits are maximised by using size, service frequency and price as control variables. Based on empirical data they find that both size and frequency vary approximately with the square root of demand. This underlines that also with much more complex models the square root principle seems to make sense.

Jansson (1993) formulates a model for a welfare maximising public transport authority that considers price and frequency. Two forms of schedule delay are distinguished: one where frequencies are so high that customers do not consult timetables when they use public transport, and another one where time tables are consulted. The two forms have rather different effects on schedule delay costs and hence may lead to local optima in the frequency choice problem.

The literature surveyed above focuses on bus transport. It is however equally relevant for rail transport. Given the nature of rail operations the number of constraints in the planning of network structures, timetables, vehicle capacities and crew and vehicle schedules tend to be more complex compared with those of bus companies (Daduna and Wren, 1988, Daduna et al., 1995). This may be an explanation why in the rail sector stylised models in terms of frequency and vehicle size only are not very common. Nevertheless, it may be argued, that although models in the tradition discussed above give a simplified picture of the optimisation of rail operations, they are useful to analyse the basic trade-offs faced by the planners.

3. CHOICE OF FREQUENCY AND VEHICLE SIZE BY PUBLIC TRANSPORT OPERATORS

Based on the approaches mentioned above we consider four possible cases for the optimisation of vehicle size and service frequency by public transport operators (see Table 1). First, we consider the dimension of elastic versus inelastic demand. Second, we distinguish operators according to the objective they maximise: profits versus social surplus. This allows us to investigate the implications of these various dimensions.

	maximise social welfare	Maximise profits
inelastic demand	I	II
elastic demand	III	IV

Table 1. Four choice contexts of frequency and vehicle size by public transport companies.

Let demand for trips depend on generalised costs GC, where GC depends on the fare p, rescheduling costs F and other travel cost components tc (cost of in vehicle time plus costs of travelling to and from railway station)

$$GC = p + tc + a/F$$

Demand also depends on other factors such as income, supply of competing modes, which are incorporated in a factor A. Thus, the demand for trips is:

$$Q = A.[p+ tc + a/F]^z \qquad (7)$$

where z is the generalised cost elasticity of demand (z<0).

The costs of the production of transport services consist of various elements. Per passenger the costs of ticket counters, cleaning, and other personnel are equal to u. Another part depends on frequency; examples are the costs of drivers, and the cost of infrastructure use). Energy costs are proportional to frequency, but are subject to economies of scale: large vehicles are more energy efficient per seat than small vehicles. This leads to a formulation of energy costs such as $C_{energy} = wFS^b$, where b=1. In a similar way the capital costs of the driving stock are assumed to display scale economies, large vehicles are cheaper per seat than short vehicles: $C_{driving\ stock} = rFS^c$. Thus, total costs are equal to:

$$C_{operator} = uQ + vF + wFS^b + rFS^c \qquad (8)$$

The costs of travellers have already been formulated in (2)

We assume that there are no restrictions on the size of the vehicles that the operator can purchase, neither are there restrictions on the frequency (thus, there is no public service obligation to have at least one train per hour, say). We impose the restriction that that total capacity F.S should be at least equal to demand: F.S = Q. Note that if some combination of F and S is found that equals Q, an increase of S, keeping F constant would never be beneficial for travellers (their utility does not depend on S), whereas for the operator it would lead to higher costs. Hence in the present formulation we will always find that

$$F.S = Q \qquad (9)$$

so that the occupancy rate Q/[F.S] always equals 1 in these models[3].
We will now discuss the four cases separately.

3.1 Maximise social welfare, inelastic demand.

This case can be considered as the direct extension of the Mohring model to the joint determination of frequency and size.

The optimal frequency F_I^* is the solution of:

$$F = [a.Q_0]^{0.5}/[v+w(1-b)Q_0{}^bF^{-b} + r(1-c)Q_0{}^cF^{-c}]^{0.5} \qquad (10)$$

The optimal vehicle size follows as $S_I^* = Q_0/F_I^*$
The square root form can still be recognised in the equations, but note that F appears both at the right hand side and the left hand side of the equation and that there is no analytical solution for it. In the extreme case that scale economies in energy use and costs of vehicle stock are absent (b=c=1) we find again the exact square root formula where the only cost component that affects frequency is v, leading to a relatively high frequency.

When the scale economies in the costs of energy and the driving stock are relatively high (b and c are close to 0), it pays the operator to employ relatively large vehicles so that frequencies will be low.

Note that since $F.S=Q_0$, the elasticities of F and S with respect to Q_0 add to 1 by definition. The introduction of the size related elements in the cost function makes the supply of frequency slightly less elastic to changes in demand than in the base case. Figure 2 illustrates the relationship between demand and the optimal frequency level. The elasticity implied by the figure is about 0.4 [4]. The response of vehicle size with respect to travel demand is illustrated in the lower part of the figure. Thus we arrive at the conclusion that the introduction of the economies of scale in the cost function makes the frequency response with respect to changes in traveller demand slightly less elastic. The difference with the standard square root result is not large, however.

Figure 2. The relationship between travel demand, frequency and vehicle size(case I).

3.2 Maximise profits, inelastic demand.

The case of maximisation of profits under the assumption of inelastic demand is a trivial one. In this case the operator would minimise costs without paying attention to the interest of the travellers since consumer loyalty is not affected. This would lead to a choice of a very low frequency and a very large vehicle size.

3.3 Maximise social welfare, elastic demand.

The maximisation of social surplus with elastic demand means that the operator's objective can no longer be formulated in terms of costs, but that it should be in terms of consumer welfare (measured by means of consumer surplus) and costs. The inverse demand function is $GC = (Q/A)^{1/z}$. Thus, consumer benefits equal

$$CB = \int_0^Q [q/A]^{1/z}\, dq$$

In the case of welfare maximisation the objective is to maximise consumer benefits minus total costs of producing and consuming the services:

$$\text{Welfare} = \int_0^Q [q/A]^{1/z}\, dq \; - [tc+a/F]Q - uQ - vF - wFS^b - rFS^c$$

Note that with the given model formulation the size S of the trains equals Q/F since given any Q and F adding capacity by means of larger trains only leads to higher costs without any benefits for the operator or the consumer. Maximisation of total welfare with respect to price and frequency yields the following first order conditions:

$$(p-u) = w.b.\ A^{b-1}.(p+tc+a/F)^{b.z-z}\ .\ F^{1-b} + r.c.\ A^{c-1}.(p+tc+a/F)^{c.z-z}\ .\ F^{1-c} \qquad (11)$$

and:

$$(p-u).A.z\ .(p+tc+a/F)^{z-1}.(-a/F^2) = v -[a/F^2].A.[p+tc+a/F]^z$$

$$+ w.b.z.A^b\ (p+tc+a/F)^{b.z-1}.(-a/F^2).F^{1-b} + w.A^b.(p+tc+a/F)^{b.z}.(1-b).F^{-b}$$

$$+ r.c.z.A^c\ (p+tc+a/F)^{c.z-1}.(-a/F^2).F^{1-c} + r.A^c.(p+tc+a/F)^{c.z}.(1-c).F^{-c} \qquad (12)$$

Results for the optimal frequencies as a function of autonomous demand (A) are given in Figure 3. It appears that, although a rather complex model is used the elasticity implied by this figure is slightly above 0.5. Thus, the introduction of endogenous demand (being dependent on frequency) does not lead to elasticities that are very different from 0.5 as indicated in the original Mohring model. Note that an increase in exogenous demand A has two effects total demand: a direct and an indirect one. The direct effect is proportional according to equation (7). The indirect effect means that generalised costs decrease when frequency increases. Thus when exogenous demand A increases with 1% total demand Q will increase with more than 1 %. Since we assume in this section that total capacity F.S is equal to total demand Q, the conclusion is that the elasticity of frequency and size with respect to A are (slightly) higher than they are with respect to Q

Figure 3. Relationship between travel demand and vehicle size under elastic demand, welfare maximisation (III) and profit maximisation (IV).

3.4 Maximise profits, elastic demand.

Profits of the monopolist are equal to:

$$Z = (p-u)Q -v.F-w.F.S^b -r.F.S^c$$

Maximisation of profits with respect to price and frequency yields:

$$[p+tc+a/F]/z + (p-u) =$$

$$w.b.\ A^{b-1}.(p+tc+a/F)^{b.z-z}\ .\ F^{1-b} + r.c.\ A^{c-1}.(p+tc+a/F)^{c.z-z}\ .\ F^{1-c} \qquad (13)$$

and

$$(p-u).A.z\ .(p+tc+a/F)^{z-1}.(-a/F^2) =$$

$$v + w.b.z.A^b.(p+tc+a/F)^{b.z-1}.(-a/F^2).F^{1-b} + w.A^b.(p+tc+a/F)^{b.z}.(1-b).F^{-b}$$

$$+ r.c.z.A^c.(p+tc+a/F)^{c.z-1}.(-a/F^2).F^{1-c} + r.A^c.(p+tc+a/F)^{c.z}.(1-c).F^{-c} \qquad (14)$$

The difference between (12) and (14) is that consumer's surplus is not taken into account by the profit maximising operator. Ignoring consumer benefits of high frequencies will of course lead to a rearrangement of capacity F.S towards larger vehicles S and lower frequencies. Note also that when z tends to zero in (14), also frequency will tend to zero, so that again the result of case II is found. Figure 3 clearly demonstrates that in the case of profit maximisation the level of frequency is clearly lower compared with the social surplus maximisation. Thus a profit oriented firm will run less (but longer) trains than a welfare maximising firm. Note, however, that there is not a large difference in terms of the elasticity of frequency with respect to changes in travel demand; in both parts of Figure 3 the elasticity is slightly higher than 0.5.

4 OCCUPANCY RATES IN PUBLIC TRANSPORT; WHY ARE THEY LOW?

Occupancy rates tend to be low in public transport in many industrialised countries. For example the average occupancy rate of railway services in the Netherlands (peak plus off-peak) is about 40%. This means that especially outside the peaks occupancy rates will be low. Here we list a number of possible reasons. We start with reasons that lead to low occupancy rates (4.1-4.4); in 4.5-4.7 we discuss some reasons that may lead to low occupancy rates, but not necessarily so. Finally, in 4.8 factors leading to high occupancy rates are mentioned.

4.1 Public service obligations:

-Public service obligations during periods with low demand.
Busses and trains tend to be rather empty early in the morning and late in the evening, as a result of the objective to supply services during virtually all relevant hours of the day.
-Public service obligations in zones with low demand.
An example is the operation of public transport services in sparsely populated areas.
-Anticipation of future demand
Some public transport companies have the policy to serve recently developed neighbourhoods that are not yet completed to offer services for residents from the very first day that they start living there. The background is the fear that residents build up a 'car orientation' so that public transport will not be chosen when it is offered at a later stage.

4.2 Spatial imbalances:

-The back-haul problem.
Spatial imbalances of demand for transport lead to the situation that when trains are needed to bring passengers from A to B during the morning peak, these trains will meet little demand for the return trip during the same peak. Thus when they make the return trip to A they may be rather empty.
-Different levels of demand at different parts of the railway line.
Demand for transport services often varies substantially between different parts of a railway line. When costs of adjusting frequencies and vehicle size are high (see below under inflexibility), this may imply excess supply of capacity at a part of the line.

4.3 Convenience for travellers:

-Regularity of time tables.
Some public transport companies have a policy of using quite regular timetables (for example one bus per hour). This implies that when during part of the day there is less demand the company will not respond by reducing its frequency. The advantage of regular timetables is that they reduce costs of information search by travellers.
-Comfort considerations.

Moderate levels of discomfort due to crowding in public transport already are found when the occupancy rate is above 50% because travellers have problems with space (also for luggage). Thus, suppliers of public transport may intentionally plan their activities in such a way that occupancy rates will stay considerably below 100%

-Search for a seat.

When trains are long and almost full, it may take a long time for passengers to find a seat, when they happen to have entered it at the 'wrong' side. This may lead public transport companies to add some capacity to save the passenger the effort of finding a seat.

4.4 Uncertainties:

-Stochastic demand.

Most public transport does not require the advance reservation of a seat. Hence suppliers face the problem of planning capacity without knowing exactly the number of passengers. The objective to offer a guaranteed space implies that a certain level of excess capacity has to be planned above the average level of demand.

-Uncertainty about the mixture of passengers in the case of heterogeneous products.

Trains usually offer four types of seats: smoking versus no-smoking and first class versus second class. This is an extra reason of low occupancy rates, because public transport companies tend to follow the policy that for each of the segments there should be sufficient capacity. Thus even when total demand would be known, but the shares of the various segments are not, there would be a tendency to supply more seats than necessary.

-Unreliability.

A certain part of the vehicles of public transport firms will not run for various reasons such as technical difficulties, personnel problems or otherwise. Therefore public transport companies may anticipate this problem by planning the deployment of more vehicles than actually needed, implying that on certain days there will be more capacity than actually needed.

In addition to the above reasons that will lead to low occupancy rates, there are a number of other reasons that *may* lead to low occupancy rates, but *not necessarily* so:

4.5 Inflexibility:

-Inflexibility of vehicle size.

Once a public transport firm has decided to invest in vehicles of a certain size, it cannot easily change its size when it appears that the vehicles have to be used for a market with less demand. Transport vehicles are characterised by indivisibilities.

-Inflexibility to adjust the composition of vehicles

It is often impossible to vary the shares of the various types of seats (smoking versus non-smoking, first class versus second class) in vehicles according to short term fluctuations in demand.

-Operational costs of changing the capacity of vehicles.

One might argue that with trains it is easier to adjust capacities than with busses, since carriages may be coupled and de-coupled. The problem is that coupling and de-coupling takes time and that the costs may be so high that it is more efficient to accept low or high occupancy rates on part of the trajectory.

-Inflexibility of time tables.

Time tables are usually fixed for periods of 6 or 12 months. If there is an unexpected increase or decrease of demand, during such a period deviating occupancy rates may emerge. There may also be predictable fluctuations in demand related to for example holidays where the operator does not adjust the timetable.

-Inflexibility of fleet size.

Changing the size of a fleet takes time, especially adding new vehicles may last some years. When demand is growing, this may lead to a high occupancy rate.

4.6 Network interdependencies:

Public transport companies usually try to operate integrated networks with possibilities to change from one vehicle to the other. This implies that frequencies of service have to be co-ordinated on various lines so that on some lines frequencies will be deviating from those actually warranted by demand on the line itself.

4.7 X-inefficiencies:

Ignorance and bad planning of activities may lead to a lack of balance between demand and supply.

Finally there are some reasons leading to the possibility of extremely *high* occupancy rates.

4.8 Capacity limits:

Physical capacity.

Every transport network has its capacity limits in terms of frequency (for example: 30 metro's per hour on a certain line), or size (for example: the length of platforms only allows trains with at most 5 carriages). Thus, link related capacity constraints may put a maximum on the frequency, whereas node related capacity constraints (length of platforms) has implications for the size of trains.

High costs of adding capacity.

Adding capacity to meet peak demand leads to high marginal costs. When public transport companies cannot charge users during the peaks for these costs, they will be reluctant to supply these additional services, so that during peaks extreme occupancy rates may emerge.

Unexpected rapid growth of demand with a short run capacity level.

When capacity as determined by the vehicle stock and personnel is given in the short run, a sudden increase in demand cannot be accommodated, leading to high occupancy rates.

5. EMPIRICAL PATTERNS OF SERVICE FREQUENCIES AND VEHICLE SIZE.

The above computations have been based on stylised facts about cost structures and demand functions in rail transport. One may wonder how the actual planning of frequencies and vehicle size compares with these theoretical results on optimal planning of operations.
To address this issue we have considered a sample of rail connections within The Netherlands. For 82 connections (46 stopping train and 36 intercity train connections) we obtained data of the number of passengers, the daily frequency and the number of seats. During the off-peak period the average occupancy rate is 51% in stopping trains and 38% in intercity trains. A nice illustration of the problem of spatial imbalances leading to different levels of demand at different parts of the railway line is given in Table 3. Near the large city with railway stations A and B occupancy rates are high, but at the end of the line the occupancy rate is almost negligible. The number of seats supplied by the railway company, and the service frequency is the same at all places, which is an indication of the inflexibilities involved as well as the public service obligations at stake.

Connection	Occupancy rate (%)
AB	62
BC	86

CD	49
DE	38
EF	29
FG	25
GH	14
HI	4

Source: NSR.

Table 2. Occupancy rates of trains (off-peak) on various parts of a railway line in the Northern part of the Netherlands (1999)

In figure 3 we present the plots of service frequency and number of seats per train for the stopping trains and intercity trains during off-peak periods. It appears that in our sample the maximum frequency observed is about 60 trains per day (off peak). Based on an operational period of about 14 off-peak hours per day this means that there are 4 or 5 trains per hour. The basic pattern of 1, 2 or 4 trains per hour is clearly visible in the figure. Variations may occur since in the early and late hours frequency may be lower. In addition, international trains may lead to some irregularity in the time-tables. Low frequencies are clearly overrepresented in the stopping train segment. Most intercity trains have a frequency of 2 per hour. Also in terms of size there is a substantial difference between intercity trains and stopping trains, the maximum size of the stopping trains considered being about 200 seats, and that of intercity trains being 700 seats.

Figure 3 Service frequencies and number of seats per train for a sample of stopping trains and intercity trains in the Netherlands (off-peak, 1999).

We consider the empirical relationship between frequency and size supplied on the one hand and the number of passengers on the other:

$$\begin{aligned} \ln F_i &= a_0 + a_1\, Dum_{int,i} + b_0 \ln Q_i + e_i \\ \ln S_i &= c_0 + c_1\, Dum_{int,i} + d_0 \ln Q_i + e_i \end{aligned} \qquad (18)$$

Where $Dum_{int,i} = 1$ when the link considered (i) is an intercity service, otherwise $Dum_{int,i} = 0$. Estimation results of these relationships are reported in Table 4.

(a) separate analysis for stopping trains and intercity trains.

	frequency; stopping train	frequency; intercity trains	size; stopping train	size; intercity train	# of seats; stopping train	# of seats; intercity trains
constant	0.442 (0.352)	0.08867 (0.369)	3.924 (0.266)	4.661 (0.310)	4.365 (0.542)	4.750 (0.376)
Dum_{int}	-	-	-	-	-	-
ln Q	0.387	0.396	0.138	0.183	0.525	0.578

	(0.047)	(0.041)	(0.035)	(0.035)	(0.072)	(0.042)
R^2	0.609	0.730	0.256	0.449	0.547	0.847

(b) joint analysis for stopping trains and intercity trains.

	frequency; all trains	size; all trains	# of seats; all trains
constant	0.423	3.826	4.249
	(0.248)	(0.193)	(0.360)
Dum_{int}	-0.282	1.118	0.836
	(0.073)	(0.057)	(0.107)
ln Q	0.390	0.151	0.541
	(0.033)	(0.026)	(0.048)
R^2	0.678	0.926	0.870

Table 3. OLS estimation results of dependence of service frequency and vehicle size on passenger demand.

The estimation results indicate that a 1% increase in passenger demand leads to a 0.4% increase in frequency and a 0.15% increase in the number of seats during off peak periods. Thus, when levels of demand vary, the railway company tends to respond more strongly in terms of frequency than in terms of size. Note also that the overall elasticity of the railway company in terms of total capacity (F.S) with respect to passenger demand is about 0.55. This is clearly below the elasticity of 1 implied by the models in section 3. This result implies that occupancy rates tend to increase with passenger demand. Note that this result is contrary to the result in section 5 where frequency is found to be more responsive. This calls for a deeper investigation of the difference between the empirical and the theoretical analysis in this paper.

An implicit assumption of the above analysis is that the correlation between the error term e and the independent variable Q is zero in equations (18). In the present context it is probable, however, that the two are correlated. To correct for this problem 2-SLS has been applied. This results in an elasticity that is somewhat lower (0.34 in stead of 0.39).

It is interesting to confront these results with the planning routines of the railway company concerned as explained by railway officials. Frequencies are at least one per hour as a consequence of the social function of railway services. In situations of higher demand, the frequency is two per hour, or four per hour. The possibility of three per hour is not employed because it would make connections on various lines incompatible. In some cases there may be deviating service frequencies, for example when international trains make use of the lines. Concerning the number of seats the planning of capacity takes place in such a way that there must be a seat for every passenger during both peak and off-peak. The only exception is that in stopping trains an occupancy rate is allowed of 1.15 during the peak. The background is that stopping trains are usually used for rather short trips, so that the loss of comfort is considered to be acceptable here.

The phenomenon of spatial imbalances usually does not lead to reductions of the number of seats at the low demand segments of the line. The reason is that uncoupling and coupling take time and may aggravate reliability problems. The consequence is that often a carriage remains in use during the whole day in the whole network when it is only needed at some places during the peak. This obviously leads to low occupancy rates during the off-peak period.

6 Discussion and conclusion.

The above results have important implications for the marginal costs -and more in particular the marginal external costs- of rail transport. The external costs of various transport modes have been estimated in several studies (see for example CE, 1999, Mayeres and Proost, 1996, Kageson, 1998). A common result is that the external costs of the car per passengerkm are not as high as is often thought when they are compared with the external costs in public transport (bus, train). The popular notion that pollution, noise, accidents and congestion per passengerkm is much higher for car users than for public transport travellers is not always supported by the empirical results.

One of the reasons of the disappointing performance per passengerkm is that the occupancy rates are usually rather low in public transport (see section 4). In the present paper we argue that for a proper interpretation of these results the distinction between average and marginal costs is of crucial importance. Low occupancy rates in public transport indeed lead to a disappointing performance in terms of the average environmental cost per passenger. However, the low occupancy rates imply that additional passengers can easily be accommodated so that there is little need to increase the volume of vehicle kilometres. The conclusion is that when there is no proportional response from the public transport operator, marginal environmental costs of public transport may be substantially lower than the average external costs. This is indeed confirmed by our empirical analysis: an increase in demand leads to a less than proportional increase in capacity and hence in environmental burden.
In addition it appears that the operator tends to be more responsive to changes in demand by adjusting the frequency, than in terms of vehicle size (number of seats). The values of the estimated elasticities imply that the marginal external costs are about 50% lower than the average costs of railway passengers.

We discussed a large number of reasons for the low occupancy rates in rail transport. These issues have received relatively little attention in economic analysis of capacity planning by public transport operators. The low occupancy rates deserve more attention in future research on capacity planning of transport firms and its cost implications. Related issues for further research concern the use of the frequency instrument to stimulate demand and a more explicit treatment of peak versus off-peak conditions.

Acknowledgement.

The author thanks NSR for providing data on rail operations in The Netherlands. Edwin Hinloopen and Harry van Ooststroom gave helpful comments.

REFERENCES

CE (1999) Efficiente prijzen voor het verkeer, Delft.

Daduna, J.R. and A. Wren (eds.) (1988) *Computer aided transit scheduling,* Springer, Heidelberg

Daduna, J.R., I. Branco and J.M.P. Paixao (eds.) (1995) *Computer aided transit scheduling, Springer,* Heidelberg.

Glaister, S. (1986) Bus deregulation, competition and vehicle size, *Journal of Transport Economics and Policy,* vol. 20, pp. 217-244.

Gronau, R. (2000) Optimum diversity in the public transport market, *Journal of Transport Economics and Policy,* vol. 34, pp. 21-42.

Jansson, J.O. (1980) A simple bus line model for optimisation of service frequency and bus size, *Journal of Transport Economics and Policy,* vol. 14, pp.53-80.

Jansson, K, (1993) Optimal public transport price and service frequency, *Journal of Transport Economics and Policy*, vol. 27, pp. 33-50.

Kageson, P. (1993) *Getting the prices right*, T&E Brussels.

Mayeres, I., S. Ochelen and S. Proost (1996) The marginal external costs of urban transport, *Transportation Research D,* vol. 1, 1996, pp. 111-130.

Mohring, H., *Transportation Economics*, Ballinger, Cambridge, 1976.

Van den Brink, R., and M. Gijsen (2000) De milieubelasting van electrisch personenvervoer, RIVM, Bilthoven.

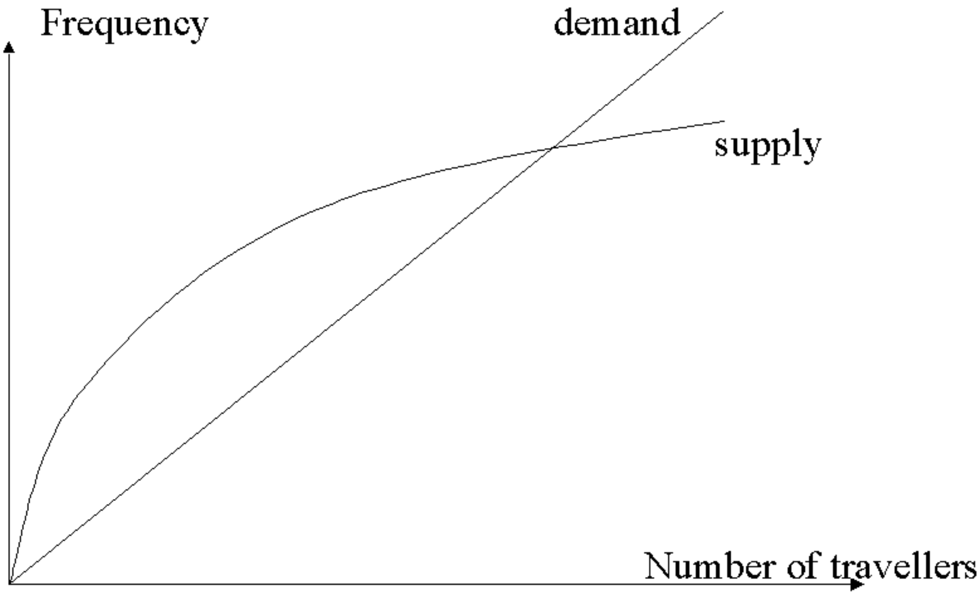
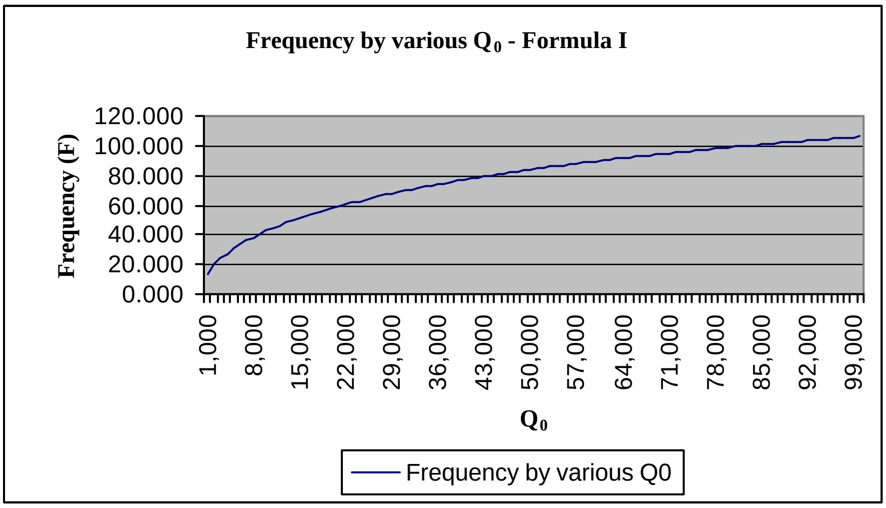


Figure 1. The square root formula for frequency



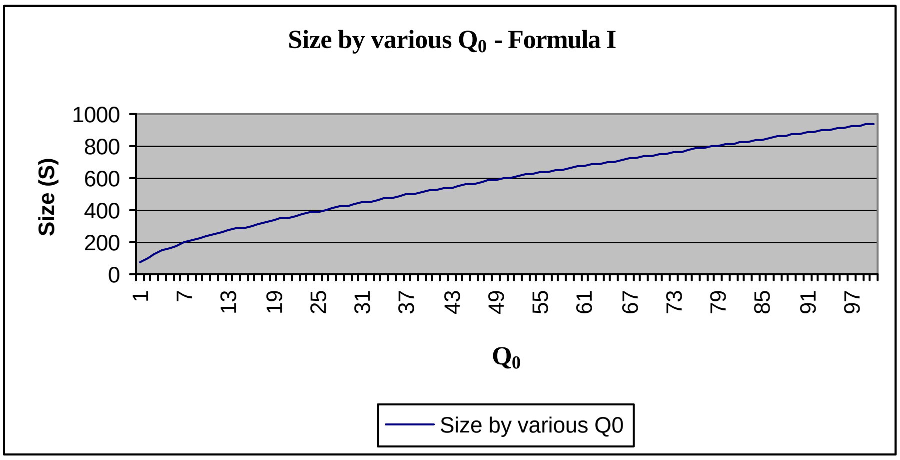
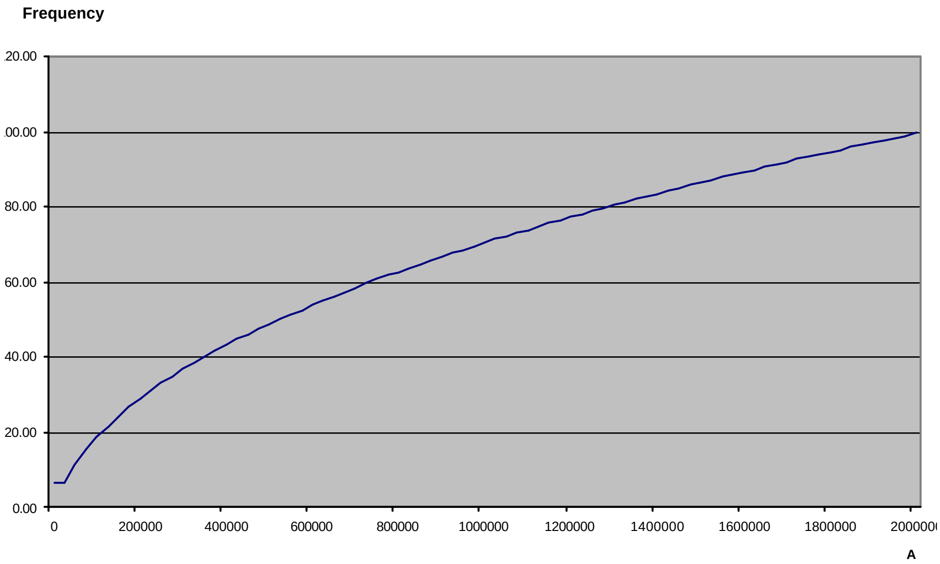
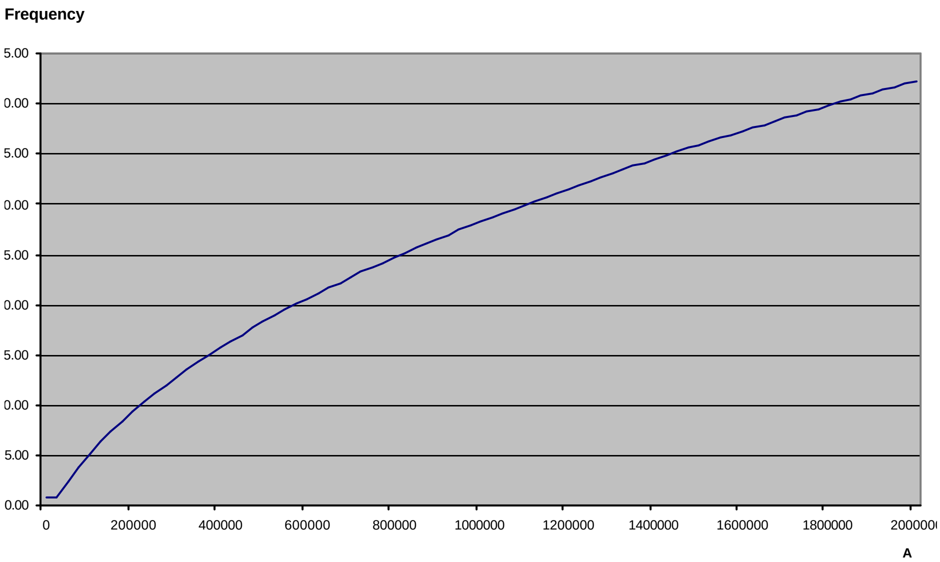


Figure 2. The relationship between travel demand, optimal frequency and vehicle size for a welfare optimising transport company under inelastic demand.



(III: welfare maximisation)



(IV: profit maximisation)

Figure 3. Relationship between autonomous travel demand A and frequency under elastic demand, welfare maximisation (III) and profit maximisation (IV).

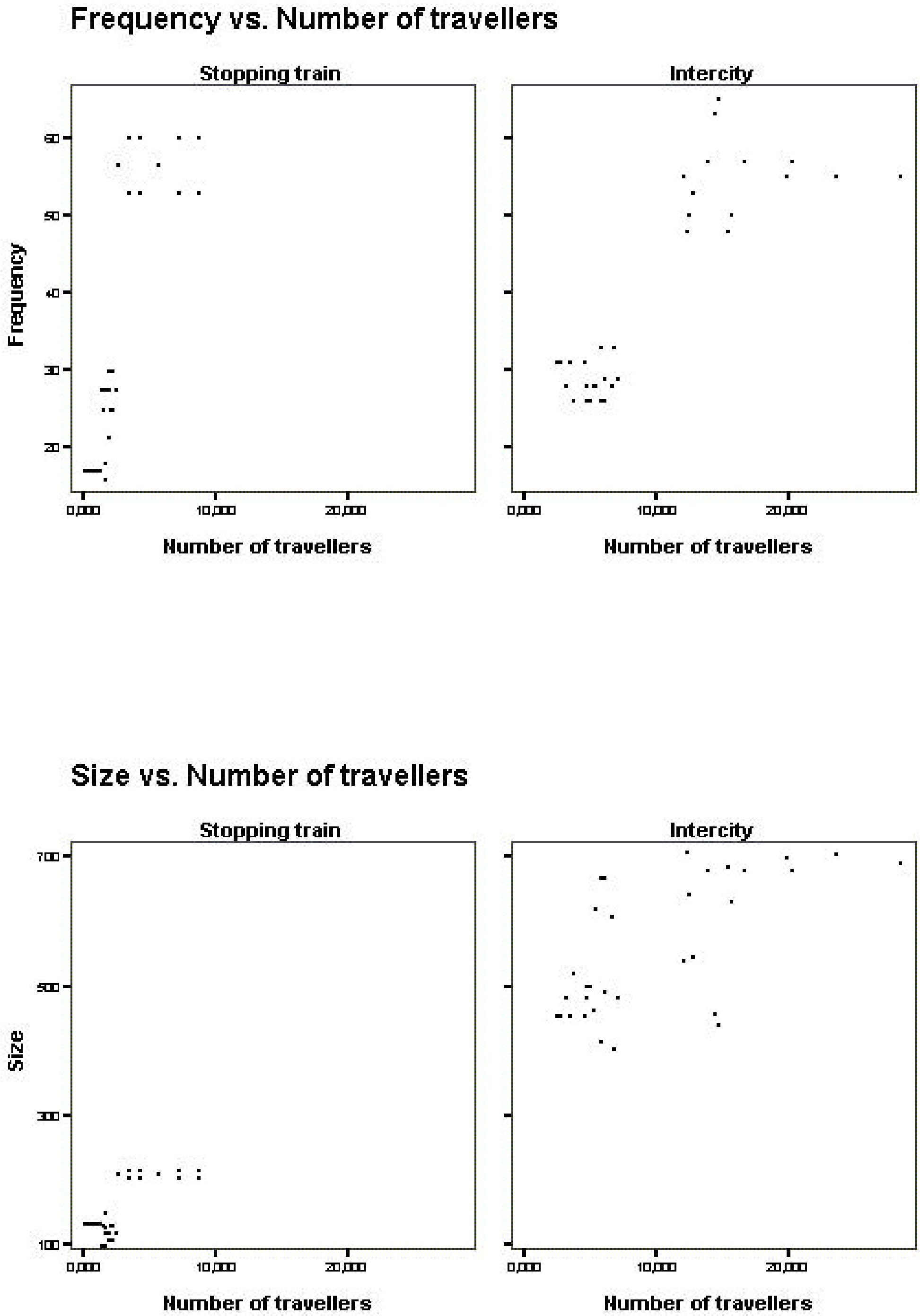


Fig. 4. Service frequencies and number of seats per train for a sample of stopping trains and intercity trains in the Netherlands (off-peak, 1999)

Endnotes.

[1] Note that we do not take into account delays related to boarding and alighting in this formulation

[2] Whether it will be below 1 for low values of Q depends on the cost parameters.

[3] In the empirical part of this paper we will address the issue of occupancy rates that are lower than 100%.

[4] The figure is based on the following combination of parameters: Q_0=20000; A=10.000; z=0.2; p=8; u=2; v=1000; w=0.8; r=3.2; b=0.7; c=0.9; a=200.