wctr'04
4-8 july 2004 / istanbul-turkey

SECOND BEST DECISION MAKING OF RAILWAY OPERATORS: HOW TO FIX FARES, FREQUENCY AND VEHICLE SIZE

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

Abstract

Railway networks are characterised by variations in demand on different links. Optimal strategies therefore call for a differentiated treatment of fares, frequencies and vehicle sizes in various links. However, for various reasons, railway operators may apply uniform levels for these decision variables. In this paper we investigate the welfare losses implied by uniform setting of fares per km, frequencies and vehicle sizes. We demonstrate that the largest welfare loss results when frequencies are made uniform across links. Welfare losses due to making vehicle size and price per km uniform across links are smaller.

Keywords: Railways; Fares; Second best; Asymmetric demand
Topic Area: D3 Integrated Supply / Demand Modelling

1. Introduction

Suppliers 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 railway operations and pay special attention to three of these instruments: choice of frequency of service, vehicle size and price. We will investigate the potential contribution of these instruments to achieving a profit or welfare maximum.

An important feature of public transport is that on various segments within networks there are substantial differences in demand. For example, the closer one gets to a large city, the higher traveller volumes become. In addition there are large differences in demand according to the direction of travel. Table 1 gives an illustration for a train service between A, B and C, where C is a large city. In the morning peak demand for train services is assumed to be high for the BC part, 50% lower for AB, and lower again in the opposite direction. As demonstrated in the table a policy to accommodate demand at the BC section leads to a rather low average occupancy rate of below 50% even during the morning peak.

Table 1. Directional and spatial asymmetries in travel demand; implications for occupancy rates.

Link	Demand during morning peak; passengerkms	Capacity: seatkms	Occupancy rate (%)
AB	500	1,000	50
BC	1,000	1,000	100
CB	250	1,000	25
BA	125	1,000	12.5
Total	1,875	4,000	46.8

The directional imbalances are usually costly to solve in railway operations. The back-haul problem is essentially a matter of joint production: if services are produced in one direction, there will also be services in the opposite direction. It will appear that opportunities to address the demand imbalances by adjustments in capacity –size, frequency- are limited[1]. It is here that pricing measures –implying direction dependent pricing- may be expected to be of vital importance (see for example Rietveld and Roson, 2002).

As Table 1 indicates, in addition to direction dependent imbalances in travel demand, there are also imbalances given differences in demand on various segments (AB versus BC). Here variations in frequency and size are easier to apply. However, railway operators often have considerable reservations against this. For example, variations in train size may imply costs of coupling and decoupling. Variations in frequencies may be difficult to implement within given structures for timetables. For example, frequencies are usually set according to fixed rules such as 2 per hour, or 4 per hour. Then a frequency of 3 per hour on a certain link may lead to long waiting times for transfer passengers when in the rest of the network the frequency equals 4. Further, differentiation in prices may confuse customers and may stimulate travellers to make detours. *The question addressed in this paper concerns the potential contribution of variations in frequency, vehicle size and fares in order to achieve a profit or welfare optimum for a railway network where demand varies between links. More in particular the performance of second best strategies will be investigated, i.e., what are the losses in terms of welfare and profits when prices, frequencies or vehicle sizes are kept uniform.* This is an important theme that has received little systematic attention in railway research.

In order to address these questions we will first give a short review of the literature on behaviour of public transport operators (section 2). In section 3 we present results for a simple model based on inelastic demand. This is followed by section 4 where the case of elastic travel demand is considered.

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

$$C_{operator} = u.Q+v.F \tag{1}$$

where u is the marginal cost per passenger and v is the cost of an extra vehicle used to serve passengers[2]. In addition to the costs experienced by the operator there are also costs for the passengers related to waiting time, schedule delay, the fare p and other travel cost components (cost of vehicle time plus costs of travelling to and from railway station). 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 timetable 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 $a_1 \geq 1$). 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:

[1] We do not go into details here on strategies that are sometimes available such as letting trains wait near the work location after the morning peak until the start of the afternoon peak, or using the excess capacity in other parts of the network.

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

$$C_{traveller} = [p + tc + a/F]Q \tag{2}$$

Minimising the sum of total costs of company and travellers

$$C = C_{operator} + C_{traveller} = u.Q + v.F + [p + tc + a/F]Q \tag{3}$$

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 leads to the conclusion that occupancy rates will be higher in situations of high demand and one would expect a tendency of introducing larger vehicles in this case. This obviously calls for a joint analysis of choice of frequency and vehicle size by operators.

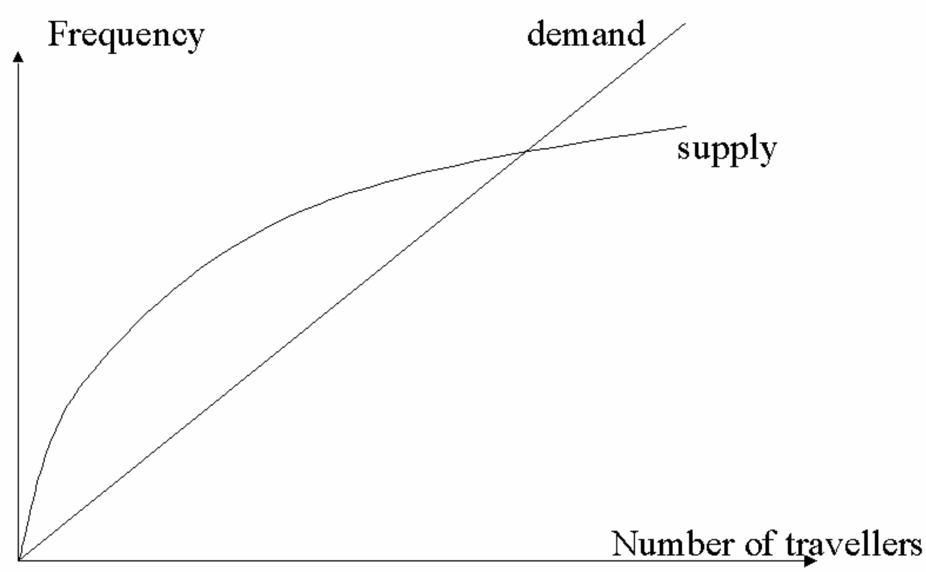

Figure 1. The relationship between capacity and travel demand according to the square root principle.

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 timetables 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. Different levels of demand at different parts of the railway line; inelastic demand

Consider now a simple network where the operator serves a line from A to C with a stop B in between. The demand on the three relevant markets is denoted as Q_{AB}, Q_{BC}, and Q_{AC}. These demands are assumed to be inelastic at this stage. We consider the case that demand is large on the AB market, and smaller at the BC and AC markets. See figure 2 for an example.

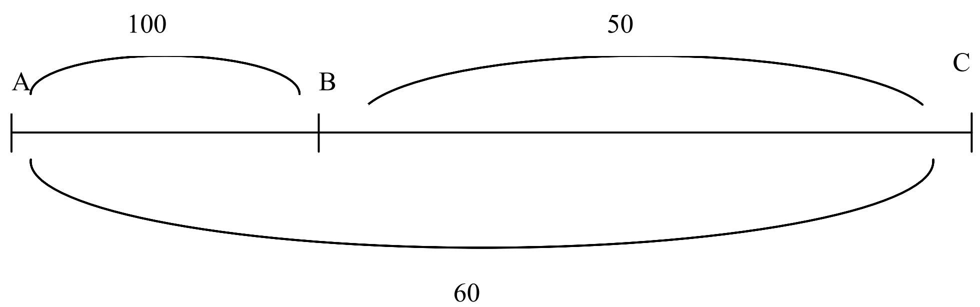

Figure 2. Railway network with three nodes and varying levels of demand.

We study various regimes for frequencies and vehicle size in terms of whether they are uniform across the whole network or not (see Table 2). Clearly, in case 1, the differences in demand levels on various parts of the network are not matched by differences in supply. The implications of the various capacity strategies for social costs will now be analysed subsequently.

Table 2. Various combinations of frequency and vehicle size according to degree of uniformity in on various links in the network

Frequency Size	Uniform	Non-uniform
Uniform	1	2
Non-uniform	3	4

Case 1: Minimisation of social costs; inelastic demand. Uniform frequency, uniform vehicle size.

As the reference case we impose the restriction that the operator applies a uniform service in terms of frequency and vehicle size on all market segments and that a uniform price per km is charged. The total number of passengers that travel between A and B equals $Q_{AB}+Q_{AC}$, and between B and C it equals $Q_{BC}+Q_{AC}$. We assume again that total capacity should be at least sufficient to meet total demand. Hence:

$$\begin{aligned} Q_{AB}+Q_{AC} &\leq F.S \\ Q_{BC}+Q_{AC} &\leq F.S \end{aligned} \tag{5}$$

Assume that demand on the AB segment is larger than on the BC segment. Then, it follows that $F.S = Q_{AB}+Q_{AC}$. We assume that demand is symmetric in both directions.

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. These costs are assumed to be proportional to distance. Energy costs per km 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 $b \leq 1$. In a similar way the capital costs of the driving stock per km are assumed to display scale economies, large vehicles are cheaper per seat than short vehicles: $C_{driving\ stock} = rFS^c$. Thus, costs of operations on a certain line of length d are equal to:

$$C_{operator} = uQ + vFd + wFS^b d + rFS^c d \tag{6}$$

Then, the total cost function of operations on the line between A and C, also including the costs of passengers (see equation (2)) is:

$$\begin{aligned} C = {} & u[Q_{AB}+ Q_{BC}+Q_{AC}] + vF(d_{AB}+d_{BC}) + wFS^b(d_{AB}+d_{BC})+ rFS^c(d_{AB}+d_{BC}) + \\ & [p_{AB} + tc_{AB} + a/F]Q_{AB} + [p_{BC} + tc_{BC} + a/F]Q_{BC} + [p_{AC} + tc_{AC} + a/F]Q_{AC} \end{aligned} \tag{7}$$

Since we assume here that demand is inelastic welfare maximisation is equivalent to cost minimisation. Another implication of inelastic demand is that price setting may be ignored. Thus, the remaining instruments are frequency F and vehicle size S. We assume that capacity F.S is set in such a way that it is just equal to demand in the busiest part of the network: S =

$[Q_{AB}+Q_{AC}]/F$. After substitution[3] of this equation in the cost function, minimisation of costs C with respect to frequency F leads to:

$$F = [a(Q_{AB}+ Q_{BC}+Q_{AC})]^{0.5} / [v(d_{AB}+d_{BC}) + w(Q_{AB}+Q_{AC})^{b}(1\text{-}b)F^{-b}(d_{AB}+d_{BC})+ r(Q_{AB}+Q_{AC})^{c}(1\text{-}c)F^{-c}(d_{AB}+d_{BC})]^{0.5} \tag{8}$$

Although this formula is already far removed from the simple square root expression in (2), the basic square root form is still clearly visible[4]. Since there is no analytical solution to (8) we use iterative methods to compute numerical values for optimal frequency . For this purpose we use stylised values for the network in terms of demand levels and distances: Q_{AB}=10,000; Q_{BC}=5,000; Q_{AC}=7,500; d_{AB}=10; d_{BC}=30; d_{AC}=40. Further, we use the following technical parameters: u=0.5; v=20; w=0.05; r=0.1; b=0.7; c=0.9. These parameters are based on studies in economies of scale in railway operations (Rietveld, 2002), and on confidential data from the National Railways on cost structures. For representative network structures the technical parameters result in cost shares for ticketing, energy, carriages/conductors and locomotives/drivers of about 12, 8, 50 and 30 percent.

Table 3. Minimisation of social costs under various assumptions of uniform frequency and size: effects on frequency, size and costs

Minimisation of social costs; inelastic demand						
			uniform freq & size	varying freq, uniform size	varying size, uniform freq	varying freq & size
F	Frequency	AB	34.29	43.41	34.86	47.24
		AC	34.29	31.01	34.86	29.69
S	Size	AB	510	403	502	285
		AC	510	403	359	421
$C_{traveller}$	Costs		192482	191349	191954	191315
$C_{operator}$			81601	73273	73920	73213
C_{tot}			274084	264622	265874	264528
OR	Occupancy	AB	100%	100%	100%	100%
	rate	BC	71%	100%	100%	100%
AC	Average cost	AB	0.15	0.16	0.15	0.17
	per passenger	BC	0.22	0.13	0.14	0.13
	km	AC	0.11	0.13	0.13	0.12
MC	Marginal cost	AB	0.33	0.04	0.18	0.16
	per passenger	BC	0.04	0.10	0.09	0.10
	km	AC	0.08	0.07	0.08	0.09

Table 3 gives the outcomes for the optimal frequency and train size. The occupancy rate in the busy part AB is 100%, whereas in the quiet part BC it is 71%. Average costs per passenger tend to be very high in the low demand segment (BC), which is of course no surprise given the low degree of capacity utilisation. Marginal costs per passengerkm are very different in the three market segments: in the busy part AB they are about 8 times as high compared with the quiet part BC. Note also the large divergence in BC between average costs (very high) and marginal costs (very low). These results are obviously caused by the excess

[3] By this substitution the costs of energy and rolling stock become dependent on demand at the AB part of the network. Travel volumes on the BC segment do not matter as long as they are smaller than at the other market.

[4] Note that under the assumption of Q_{AC}=0, Q_{AB}=Q_{BC}=Q, d_{AB}=d_{BC}=1 and 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 in equation (4).

supply of capacity implied by the restriction that in all segments there should be sufficient capacity, and that capacity is not allowed to vary between market segments.

The implications of this second best approach for the responsiveness of size and frequency with respect to changes in demand are shown in Table 4. It appears that an increase in demand on the busy segment AB of 1% leads to an increase in overall frequency and train size on this segment of .20% and .37%, respectively. An increase of demand in the quiet part leads to a small frequency increase and a small size decrease keeping total capacity constant. This means that travellers on the AB segment would benefit from an increase of demand in the quiet segment because of the increase in frequency. This is an example of a positive consumption externality in a network context. In the lower part of Table 4 the effects of demand increases on frequency and size are summarised. It appears that an increase in demand such that all market segments grow at the same rate leads to a proportional capacity response in terms of frequency and size (.46+.54=1.00). The square root principle following from the stylised case as outlined by Mohring (1976) would imply a frequency elasticity of 0.50. In the context of the more differentiated network given here we find a slightly lower responsiveness of frequency, that is, however, still very close to the value of 0.5.

Table 4. Minimisation of social costs under various assumptions of uniform frequency and size: elasticities of supply with respect to changes in demand.

	Minimisation of social costs; inelastic demand			
Elasticity	uniform freq & size	Varying freq, uniform size	varying size, uniform freq	varying freq & size
F_{AB} - Q_{AB}	0.20	0.55	0.22	0.46
F_{AC} - Q_{AB}		-0.01		0.01
S_{AB} - Q_{AB}	0.37	0.01	0.34	0.75
S_{AC} - Q_{AB}			-0.22	-0.01
C_{oper} - Q_{AB}	0.40	0.22	0.24	0.22
F_{AB} – Q_{BC}	0.12	-0.14	0.10	0.04
F_{AC} – Q_{BC}		0.25		0.17
S_{AB} – Q_{BC}	-0.12	0.14	-0.10	-0.82
S_{AC} – Q_{BC}			0.30	0.23
C_{oper} – Q_{BC}	0.08	0.20	0.19	0.20
F_{AB} – Q_{AC}	0.15	0.06	0.14	-0.02
F_{AC} – Q_{AC}		0.23		0.28
S_{AB} – Q_{AC}	0.28	0.37	0.28	0.54
S_{AC} – Q_{AC}			0.45	0.32
C_{oper} – Q_{AC}	0.30	0.35	0.34	0.35
F_{AB} – skms	0.46	0.47	0.47	0.48
F_{AC} – skms		0.47		0.46
S_{AB} – skms	0.54	0.53	0.53	0.50
S_{AC} – skms			0.53	0.53
C_{oper} – skms	0.78	0.77	0.77	0.77

Another result of our analysis is that it allows us to investigate economies of density in the railway sector. Economies of density are usually studied in the context of aggregate cost functions where network structures and asymmetries in demand at various segments are ignored. (see for example Caves et al., 1984, and Small, 1992). The present model allows the aggregation of total output (seatkms) and total costs of the operator. Then, the welfare maximisation approach adopted here leads to a total cost elasticity of about 0.78. The returns to density measure would be 1/(0.78)= 1.28, which is very close to estimates usually obtained for costs functions based on aggregate data (see for example Pels et al., 2003)

In figure 3 we illustrate the sensitivity of frequency for changes in passenger demand[5]. The figure demonstrates the rather low response of frequency with respect to demand in the busiest part of the line (AB). This underlines that in the case with spatial variations in demand, it is optimal that the operator sets frequency in a way that departs from the simple Mohring rule.

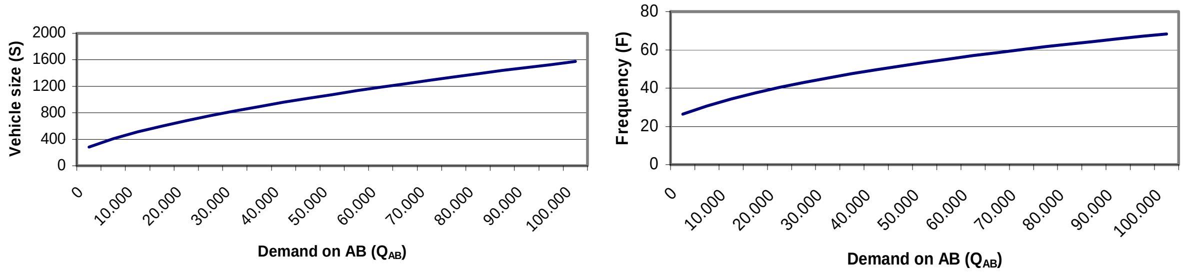

Figure 3. The relationship between travel demand, frequency, and vehicle size with different levels of demand at various parts of a railway line.

Case 2. Minimisation of social costs; inelastic demand. Varying frequency, uniform vehicle size.

We now drop the restriction that the operator applies a uniform service in terms of frequency. The frequency on the segment AB differs from the frequency on the segments BC and AC. Because the demand on AB is larger than on BC, the frequency on the first segment will be higher than on the latter. This means that trains, that arrive at B, not always continue to C, they sometimes immediately return to A. Therefore some changes in the cost function are to be applied. The costs of the public transport operator now become:

$$C_{operator} = u[Q_{AB}+Q_{BC}+Q_{AC}] + v[F_{AB}d_{AB}+F_{BC}d_{BC}] + wS^{b}[F_{AB}d_{AB}+F_{BC}d_{BC}] + rS^{c}[F_{AB}d_{AB}+F_{BC}d_{BC}] \tag{9}$$

The costs of the passengers become:

$$C_{traveller} = [p_{AB} + tc_{AB} + a/F_{AB}]Q_{AB} + [p_{BC} + tc_{BC} + a/F_{BC}]Q_{BC} + [p_{AC} + tc_{AC} + a/F_{AC}]Q_{AC} \tag{10}$$

As Table 3 shows, the result of the relaxation of the condition of equal frequency in all segments is clear: on the busy segment frequency increases, whereas on the quiet segment it decreases. From a welfare perspective the average traveller will benefit: generalised costs decrease with about 0.6%. Note, however, that this does not imply that all travellers benefit: travellers on the quiet segment are obviously better off with the high frequencies in the reference case. Note that also the costs of the railway operator decrease with about 11%. Thus, making frequency flexible has a much larger effect on operator costs than it has on traveller costs (but note that traveller costs also depend on fares, and access costs, so that indeed a substantial part of these costs cannot be influenced by the operator). Occupancy rates are now equal on all market segments, and the variation in the marginal costs of passengerkilometres among the market segments is smaller than in the reference case. Interesting enough the marginal costs are now highest in the quiet segment BC. The reason is

[5] The parameter values used are: Q_{AB}=10,000; Q_{BC}=5,000; Q_{AC}=7,500; d_{AB}=10; d_{BC}=30; d_{AC}=40; u=0.5; v=20; w=0.05; r=0.1; b=0.7; c=0.9.

given in Table 4 where it appears that frequency in the AC segment has become rather responsive to demand on BC.

Table 4 shows the responsiveness of frequency and train size with respect to demand changes. On the busy market frequency has become very responsive, whereas size has become very unresponsive here. On the other hand, size has become responsive to changes in demand on the long distance market. At the overall network level the responsiveness of size, frequency and total costs with respect to total seatkms is very close to the one in the reference case. For the cost function we observe that dropping the equal train size constraint has considerable impact on cost levels (a decrease of 11% mentioned above), but a negligible effect on economies of density estimates. Thus, constraints on railway operations may have a substantial impact on inefficiencies, while at the same time estimates of economies of density in railway operations remain unaffected.

Case 3. Minimisation of social costs; inelastic demand. Uniform frequency, varying vehicle size.

We return to the base model (uniform frequency and uniform vehicle size), but now drop the restriction of uniform vehicle size. We consider the possibility to (un)couple a railway carriage at B; in that way the vehicle size on segment AB differs from the size on BC. The costs of the public transport operator are:

$$C_{operator} = u[Q_{AB}+Q_{BC}+Q_{AC}] + vF[d_{AB}+d_{BC}] + wF[S_{AB}{}^{b}d_{AB}+S_{BC}{}^{b}d_{BC}] + rF[S_{AB}{}^{c}d_{AB}+S_{BC}{}^{c}d_{BC}] \tag{11}$$

The costs of the passengers become:

$$C_{traveller} = [p_{AB} + tc_{AB} + a/F]Q_{AB} + [p_{BC} + tc_{BC} + a/F]Q_{BC} + [p_{AC} + tc_{AC} + a/F]Q_{AC} \tag{12}$$

This alternative way of introducing flexibility by allowing varying vehicle size leads to better outcomes for both traveller costs and operator costs compared with the reference case. However, when compared with the case of flexible frequency, it is inferior. The effect on operator costs is comparable to that of variable frequency, but the development of traveller costs is not as good. This is a plausible result, since in our model formulation, changes in vehicle size only have a direct effect on operator costs whereas there is no such direct effect on traveller's utility, because nuisance due to crowding is ruled out. With changes in frequencies this is different since these have a direct impact on costs of both travellers and operator.

Case 4. Minimisation of social costs; inelastic demand. Varying frequency, varying vehicle size.

The last extension to this model is making the vehicle size segment-dependent. For example, one may allow that the vehicle size on the segment AC is larger than on the segment AB. To compose the cost function in this case, we need to define another variable, G. G denotes the frequency solely on the segment AB, that is when a train departing from A arriving at B continues to C, this train does not contribute to G_{AB}. This train contributes to G_{AC}, because this train drives solely on AC. In this way $F_{AB}=G_{AB}+G_{AC}$. By the introduction of G, we are able to determine the capacity restrictions:

$$\begin{aligned} Q_{AB}+Q_{AC} &\leq G_{AB}S_{AB}+G_{AC}S_{AC} \\ Q_{BC}+Q_{AC} &\leq G_{AC}S_{AC} \end{aligned} \tag{13}$$

The last restriction holds because G_{BC}=0. Thus, there are no trains solely on the segment BC, passengers that want to travel from B to C, travel with a train that drives between A and C. The corresponding cost function is:

$$C_{operator} = u[Q_{AB}+Q_{BC}+Q_{AC}] + v[G_{AC}d_{AC}+G_{AB}d_{AB}] + wS_{AB}^{b}G_{AB}d_{AB} + wS_{AC}^{b}G_{AC}d_{AC} + rS_{AB}^{c}G_{AB}d_{AB} + rS_{AC}^{c}G_{AC}d_{AC} \quad (14)$$

$$C_{traveller} = [p_{AB}+tc_{AB}+a/(G_{AB}+G_{AC})]Q_{AB} + [p_{BC}+tc_{BC}+a/G_{AC}]Q_{BC} + [p_{AC}+tc_{AC}+a/G_{AC}]Q_{AC} \quad (15)$$

Table 3 shows that the entirely flexible case has differentiations in size and frequency that are indeed more pronounced than the differences in the preceding alternatives. The outcome is a high frequency / small train service at the short distance and a low frequency / long train service at the long distance market. However, the differences in total costs are very small compared with the case that only frequency is flexible. This leads to the conclusion, that compared with the reference case of uniform frequency and size, varying frequency is a very well performing second best alternative to the first best solution where both size and frequency vary.

Table 4 finally gives a view on how the planning in terms of size and frequencies in the various markets is affected by changes in demand in submarkets. With uniform frequency and size, the elasticities are rather small, which is no surprise since frequency and size are assumed to serve all submarkets. In the most flexible alternative, an increase in travellers in the busiest segment (AB) has a strong effect on size (elasticity equal to 0.75), whereas size in this AB market strongly decreases with increasing demand in the other market (BC). This table shows that the responsiveness of size and frequency with respect to demand shifts in submarkets varies strongly according to the regime of fixed versus flexible size and frequency. In the lower part of Table 4 we have introduced the effects of an average increase in all submarkets. Then it appears again that the elasticities of size and frequency with respect to travel demand are both very close to 0.5 This underlines the robustness of Mohring's result derived for a simple one-line network in the context of more complex networks.

Our overall conclusion is that when demand is inelastic the second best strategy of working with non-uniform frequencies leads to outcomes for total costs that are very close to the first best strategy. Setting the range between the completely uniform service and the first best solution equal to 100 (192482-191315 for travellers and 81601-73213 for the operator) we find that the loss of keeping vehicle size uniform (case 2) is only 1 to 3% of this range. On the other hand the loss of keeping frequency uniform and only varying vehicle size is 55% for traveller costs and 8% for operator costs. Thus, compared with strategies involving different frequencies at different parts of the network, strategies dealing with non-uniform vehicle size are performing much worse.

A limitation of the present analysis is that due to the assumption of given travel demand, the price instrument cannot be considered. Therefore we shift our attention now to the case of elastic travel demand. We will address the question whether the above conclusion on the superiority of frequency as an instrument above vehicle size is still valid when demand is inelastic, and how the efficiency of these capacity oriented measures compare with the efficiency of price measures.

4. Optimisation of railway operations under elastic demand

We now consider the case that demand is elastic, so that it depends on frequency and fares. Let demand for trips depend on generalised costs GC, where GC depends on the fare p, scheduling costs that are related to frequency F and other travel cost components tc as outlined in (2). 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 on the three segments is[6]:

$$
\begin{aligned}
Q_{AB} &= A_{AB}.[p_{AB} + tc_{AB} + a/F]^{z} \\
Q_{BC} &= A_{BC}.[p_{BC} + tc_{BC} + a/F]^{z} \\
Q_{AC} &= A_{AC}.[p_{AC} + tc_{AC} + a/F]^{z}
\end{aligned}
\qquad (16)
$$

where z is the generalised cost elasticity of demand (z<0). This formulation with elastic demand means that in addition to frequency and vehicle size, also fares are considered. We consider two cases, one where the objective is the maximisation of profits, the other one is the maximisation of social welfare.

Maximisation of social welfare; elastic demand.

The maximisation of social surplus with elastic demand means that the overall objective can no longer be formulated in terms of costs, but that consumer surplus and profits have to be considered. The inverse demand function is $GC = (Q/A)^{1/z}$. Thus, consumer surplus equals

$$
\begin{aligned}
CS = & \int_0^{Q_{AB}} [q/A_{AB}]^{1/z} dq + \int_0^{Q_{BC}} [q/A_{BC}]^{1/z} dq + \int_0^{Q_{AC}} [q/A_{AC}]^{1/z} dq \\
& - [p_{AB} + tc_{AB} + a/F]Q_{AB} - [p_{BC} + tc_{BC} + a/F]Q_{BC} - [pAC + tc_{AC} + a/F]Q_{AC}
\end{aligned}
\qquad (17)
$$

In the case of welfare maximisation the objective is to maximise total surplus TS, that is consumer surplus plus profits $p_{AB}Q_{AB}+p_{AC}Q_{AC}+p_{BC}Q_{BC}$ -C:

$$
\begin{aligned}
TS = & \int_0^{Q_{AB}} [q/A_{AB}]^{1/z} dq + \int_0^{Q_{BC}} [q/A_{BC}]^{1/z} dq + \int_0^{Q_{AC}} [q/A_{AC}]^{1/z} dq - [tc_{AB} + a/F]Q_{AB} - \\
& [tc_{BC} + a/F]Q_{BC} - [tc_{AC} + a/F]Q_{AC} - u[Q_{AB} + Q_{BC} + Q_{AC}] - vF(d_{AB} + d_{BC}) - \\
& wFS^{b}(d_{AB} + d_{BC}) - rFS^{c}(d_{AB} + d_{BC})
\end{aligned}
\qquad (18)
$$

Table 5 reports the results for maximisation of social welfare. In order to prevent outcomes with negative profits, we impose the constraint that profits should be positive. Depending on the institutional setting where subsidies are given to public transport, this might be replaced by side conditions, that losses do not exceed a certain maximum level. A major lesson to be learnt from Table 5 is that flexibility of prices (column 2) is a rather ineffective way to improve welfare compared with flexibility of size or frequency.

---

[6] The parameters in this demand function have been set at the following values: A_{AB}=10,000; A_{BC}=5,000; A_{AC}=7,500; tc_{AB}=2.04; tc_{BC}=3.79; tc_{AC}=4.67; a=50, z=-1.5

Table 5. Maximisation of social welfare under various assumptions of uniform frequency, size and fare.

			Maximisation of social welfare; elastic demand							
			uniform freq, size & price	Varying price, uniform freq & size	varying freq, uniform size & price	varying freq & price, uniform size	Varying size, uniform freq & price	Varying size & price, uniform freq	varying freq & size, uniform price	varying freq, size & price
F	Frequency	AB	31.64	37.85	89.92	78.02	41.90	42.53	73.91	74.82
		AC	31.64	37.95	22.93	27.06	41.90	42.53	27.71	28.07
S	Size	AB	705	496	387	405	699	644	533	449
		BC	705	496	387	405	241	276		
		AC							341	387
p	Price	AB	2.02	3.88	1.47	1.83	1.48	1.92	1.45	1.77
		BC	6.07	1.65	4.41	3.51	4.45	3.76	4.34	3.67
		AC	8.09	4.42	5.88	4.39	5.94	4.58	5.78	4.49
Q	Demand	AB	18635	12812	30440	26071	24371	21443	29400	26332
		BC	3877	8529	4481	5417	5173	5814	4790	5333
		AC	3685	5955	4402	5550	4936	5942	4662	5526
OR	Occupancy	AB	100%	100%	100%	100%	100%	100%	100%	100%
	rate	BC	34%	77%	100%	100%	100%	100%	100%	100%
R	Revenues		90985	90116	90428	91076	88480	90295	90229	91059
C_{oper}	Costs		90985	90116	90428	91076	88480	90295	90229	91059
Z	Profits		0	0	0	0	0	0	0	0
CS	Consumer surplus		404770	425065	453007	455610	444222	445919	454403	455699
W	Welfare		404770	425065	453007	455610	444222	445919	454403	455699
Rec	Receipts per	AB	0.20	0.39	0.15	0.18	0.15	0.19	0.14	0.18
	passenger km	BC	0.20	0.05	0.15	0.12	0.15	0.13	0.14	0.12
		AC	0.20	0.11	0.15	0.11	0.15	0.11	0.14	0.11
FC	Full Costs per	AB	0.14	0.15	0.17	0.16	0.14	0.14	0.15	0.16
	passenger km	BC	0.44	0.17	0.13	0.13	0.17	0.16	0.14	0.13
		AC	0.10	0.11	0.13	0.12	0.15	0.14	0.14	0.12
NRec	Net receipts	AB	0.06	0.24	-0.02	0.02	0.01	0.05	-0.01	0.02
	per passenger	BC	-0.23	-0.11	0.02	-0.01	-0.02	-0.03	0.00	-0.01
	km	AC	0.10	0.00	0.02	-0.02	0.00	-0.02	0.01	-0.02

Compared with the reference case (column 1), flexible prices only lead to an increase of potential welfare of (425065-404770)/(455699-404770)=39.8%. As illustrated in Figure 4, this figure is 77.5% for flexible vehicle size, and for flexible frequency it is as high as 94.7%. Thus, flexible prices do help to improve the social surplus in rail transport, but changes in the supply of capacity (frequency and train size) appear to be more efficient. Another result is that when two instruments are applied in a non-uniform way, most of the potential welfare gains can be achieved. Least attractive is the combination of vehicle size and fare with a score of about 80%, the combinations of F,p and F,S get close to a score of 100%. It is interesting to note that the combination F,p performs slightly better than F,S. This reveals that the degree of overlap in the effects of F and S is larger than in the effects of F and p.

Flexible prices with uniform frequency and size imply that the per km price in the busy segment (AB) is about 7 times higher than in the quiet segment (BC) –note that we assume BC to be three times as long as AB. Thus, optimal fares indeed vary strongly in a situation with differences in demand levels, but in spite of this the price instrument is not very effective in getting close to the welfare optimum. Comparison of the first best case with the case where the price per km is uniform in the network (the two right most columns in table 5) reveals that the welfare loss of imposing a uniform price per km is very small.

There is an interesting link with the literature on road pricing. It is well known that when the degree of congestion varies between different links in the network important welfare gains can be achieved with differentiated prices. However, this result only holds true long as capacity is fixed. As demonstrated by Verhoef and Rouwendal (2004). when capacities are optimised there are no welfare gains by using differentiated road prices. This result is obtained under the assumption of constant returns to scale for road capacity. In the railway case we use cost formulations that depart from the constant returns to scale assumption. Therefore, we do find a welfare increasing effect of the introduction of non-uniform prices. However, the effect is very small, so that the main conclusion still holds, that the most efficient way to deal with variations in a network with non-uniform demand is to use differentiated frequencies. Differentiation of vehicle size achieves the second position and differentiation of fares is least attractive.

It is important to realise that different results may be obtained in the case of direction dependent variations in demand as also illustrated in Table 1. In this case, the operator has to face the issue of joint production: by producing the services between A and C, there will also be return services between C and A. When demand in both directions is strongly assymmetric, as it is in Table 1, price differentiation may become an important instrument (see Rietveld and Roson, 2002).

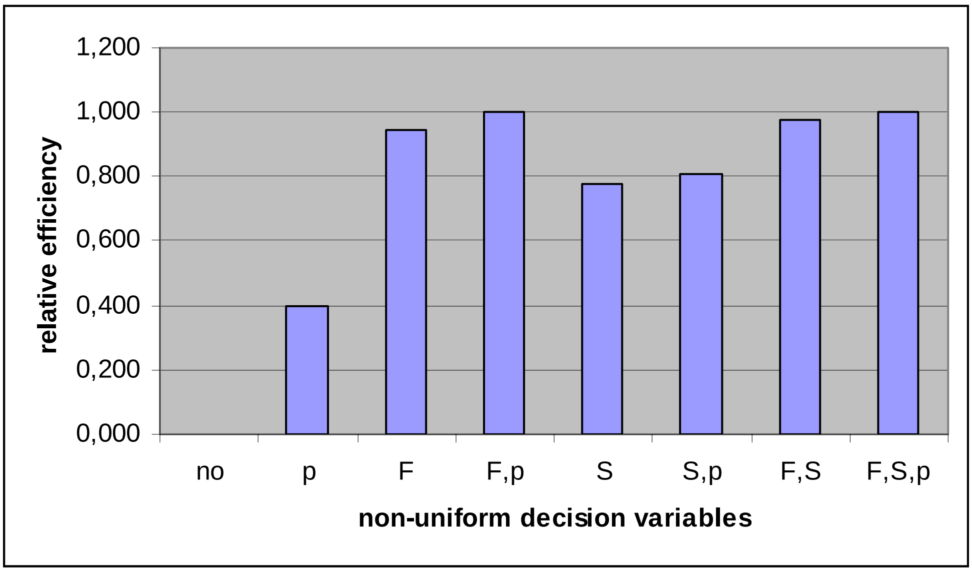

Figure 4. Relative efficiency of strategies ranging from undifferentiated frequency, vehicle size and price per km (left) to a fully differentiated strategy (right) under welfare maximisation.

Maximisation of profits; inelastic demand.
Profits of the monopolist are equal to:

$$Z = (p_{AB}-u)Q_{AB} + (p_{BC}-u)Q_{BC} + (p_{AC}-u)Q_{AC} - vF(d_{AB}+d_{BC}) - wFS^{b}(d_{AB}+d_{BC}) - rFS^{c}(d_{AB}+d_{BC}) \quad (19)$$

The maximisation of profits appears to lead to higher prices, and lower frequencies and vehicle sizes compared with maximisation of welfare. Given the purpose of our paper we focus on relative welfare performance of the various strategies, illustrated in Figure 5.

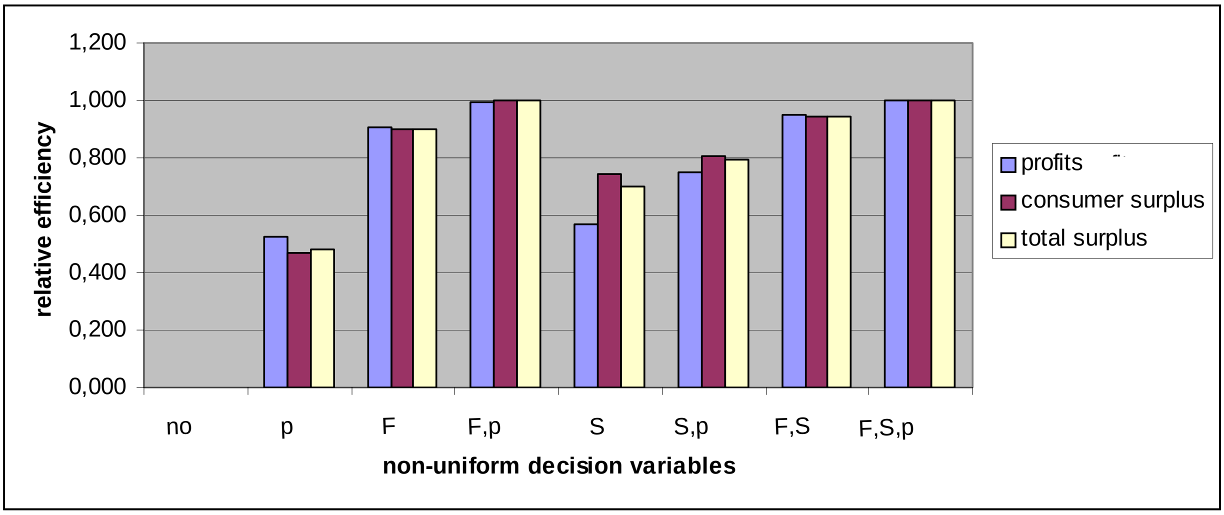

Figure 5. Relative efficiency of strategies ranging from undifferentiated frequency, vehicle size and price per km (left) to a fully differentiated strategy (right) under profit maximisation.

The relative efficiency effects of imposing uniformity constraints on fares, frequencies and vehicle size in the case of profit maximisation are similar to those of welfare maximisation, as can be observed by comparing Figures 5 and 4. In both cases the most efficient instrument to cope with variations in demand in various market segments is the use of the frequency instrument. As soon as frequencies and vehicle size have been established at their optimal levels, only small efficiency gains can be achieved by the introduction of differentiated fares per km.

5. Conclusions

There are various reasons why railway operators may wish to apply uniform frequencies, vehicle size and fares per km. These reasons include indivisibilities in frequencies and in wagon size. It is more convenient for travellers when frequencies are a fixed integer number per hour. And given network interdependencies, having a frequency of 3 per hour in a system where two per hour is the standard is not very comfortable. For vehicle size flexibility can be achieved by coupling wagons, but wagon size itself is fixed in the short run, and costs of coupling and uncoupling may be substantial. Also different prices per km may lead to resistance from travellers when these are considered as nontransparent or unfair. In the present paper we investigate the welfare consequences of the imposition of such uniformity constraints.

The use of uniform frequencies and vehicle size in railway networks with varying levels of demand lead to losses in terms of both operator costs and generalised traveller costs. When demand is inelastic the second best strategy of working with non-uniform frequencies and keeping vehicle size constant leads to outcomes for total costs that are very close to the first best strategies.

In the variants with elastic demand also the price instrument can be incorporated. We find that differentiated prices do help to increase overall efficiency of the railway system, but that the effect is small. When supply, measured in terms of vehicle size and frequency has been optimised, the potential contribution of differentiated prices is limited. Thus, only when size and frequency are not set at their optimal level, price differentiation becomes important.

An additional result is that the square root formula –derived in simple network models- still is a good approximation of the optimal response of frequency when overall demand increases in larger network models. However, when there are non-uniform changes in demand in various parts of the network quite differentiated frequency responses may be called for.

The model formulations employed here are based on detailed cost functions for individual links. They can be used to compute aggregate economies of scale in railway operations in larger networks. In our numerical example we arrive at economies of scale of about 1.28, a value that is close to what is found in the aggregate costs functions literature. It also appears that this parameter hardly depends on specific frequency or size constraints imposed on the operator. Thus, uniformity of frequency or vehicle size certainly matters for total cost levels, but not for the economies of scale parameter.

Acknowledgement.
The author thanks Stefan van Woudenberg for computational assistance.

References

Caves, D.W., Christenson, L.R. and Swanson, J.A., 1980. Productivity in US railroads, Bell Journal of Economics, 11, 65-83.

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

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

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

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

Gronau, R., 2000. Optimum diversity in the public transport market, Journal of Transport Economics and Policy, 34, 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, 14, 53-80.

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

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

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

Pels, E. and Rietveld, P., 2001. Cost functions in transport, in K.J. Button and D.A. Hensher (eds) Handbooks in Transport Volume 3. Handbook of Transport Systems and Traffic Control, Elsevier Science, Oxford, pp.321-334.

Pels, E., Daniel, V. and Rietveld, P., 2003. Economies of scale in the railway sector, Tinbergen Institute, Amsterdam, 2003

Rietveld, P. and Roson, R., 2002. Direction dependent prices in public transport: A good idea? The back haul pricing problem for a monopolistic public transport firm, Transportation, 29, 2002, 397-417

Small, K.A., 1992. Urban Transport Economics, New York, Harwood Academic Publishers.

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