Charging for Scarce Rail Capacity in Britain: A Case Study

DANIEL JOHNSON *

Institute for Transport Studies, University of Leeds

CHRIS NASH

Institute for Transport Studies, University of Leeds

Abstract

The aim of this paper is to examine the feasibility of identifying an appropriate rail scarcity charge which would make operators pay for their use of rail capacity in line with the opportunity cost of the use of these slots and to give some idea of the likely effects of such charges. The way in which we do this is to use a passenger demand forecasting model, PRAISE, to consider a situation on the East Coast Main Line which is characterized by scarce capacity and a degree of competition.

1   Background

When the government chose to privatize British Rail using a structure that separated infrastructure from operations, the result was a need to develop a method of charging for the use of rail infrastructure. Economic principles suggest that train operators will only have the correct incentives regarding speed and frequency of service and type of equipment to use, if the change in the charge levied as a result of a service level change reflects the marginal social cost of that change. In the case of rail infrastructure that marginal social cost will generally reflect wear and tear on the system (leading to changes in maintenance and renewal costs), any increased operating costs such as signaling, any external costs of accidents and environmental effects, and costs of increasing capacity utilization, including increased delays to other operators and scarcity costs (the inability of other operators to get the slots they want). If, as is likely given the economies of density in rail infrastructure, such charges are not sufficient to meet the financial requirements of the infrastructure manager, after allowing for any government contribution, then the financial requirement should be met by additional charges levied in the least distorting way possible (for example, Ramsay pricing or two-part tariffs).

The modifications brought about by the first periodic review of Railtrack's access charges (ORR, 2000) brought the structure of rail track charges in Britain much closer to the theoretical ideal. Wear and tear charges varied in fine detail according to the characteristics of the vehicle although still not representing variation by track type. A capacity charge was introduced based on the congestion cost caused by an additional train,

---

* Contact Author. Institute for Transport Studies, University of Leeds, Leeds, UK, LS2 9JT. E-mail: d.h.johnson@its.leeds.ac.uk All errors are the authors.

and it was intended that this should vary in fairly fine detail according to where and when the train operated. In practice, it is understood that the capacity charge only varies by service group and not within this by time and place because of problems with the billing software. The only charges over and above those based on marginal costs were fixed charges levied on franchised passenger operators.

Infrastructure charges should also meet the external accident and environmental costs of additional services. Given the low accident risks, and the fact that railway companies are responsible for their own insurance, it seems unlikely that the external accident costs are very large. Environmental costs have been quantified in a recent study (Sansom *et al*, 2001) although these are much smaller for rail than road, they are typically significant relative to the marginal infrastructure usage costs and should therefore be included.

One of the most difficult issues to deal with in rail infrastructure charging is that of scarce capacity. Ideally, charges would give train operators appropriate incentives to expand services only where the value of the service is at least as high as the costs it causes, and where capacity is scarce to ensure that it is used to provide the services of greatest value. This issue has become of great importance given the growth of traffic and the high costs of expansion and the consequent need to make the most effective use possible of what capacity is available (Bowker, 2002). Currently in Britain, as in most countries, decisions about the allocation of slots are taken on a purely administrative basis. This means that once an operator gains a slot, they have no incentive to seek to make more efficient use of capacity, for instance by retiming, rerouting or accelerating their train.

A capacity charge based on the cost of congestion has been introduced in Britain. However, congestion is only the appropriate capacity cost where the train in question constitutes an additional train to what would otherwise have been run; where the train in question runs instead of some other train the appropriate capacity cost is the opportunity cost of trains forced off the system by lack of capacity. Where capacity constraints bite, use of a particular slot by one train operator leads to inability of others to obtain their desired slots. Where there is a choice between operating an additional train and forcing another one off the system, obviously the course of action with the lower social cost should be taken, so it will be the lower of these two values that should be adopted.

Charging for scarce capacity would require estimation of the opportunity cost of a slot. The most attractive solution to this problem in theory is to "auction" scarce slots. There are many practical difficulties however, including the complicated ways in which slots can be put together to produce a variety of types of service and the fact that the value of a particular slot for a particular use depends on how other slots are being used (in terms of the operation of complementary or competing trains). It is also the case that the willingness to pay for the slot by the train operating company will only reflect its social value if appropriate subsidy regimes are in place to reflect the user and non user benefits of the service as discussed below. Nilsson (2002) provides a more detailed consideration of auctioning and argues that it is a feasible solution. He suggests an approach in which train operating companies bid for the timetable they want, with an indication of how their bid would be affected if the were allocated slots earlier or later than they desired. An optimization program then produces the best allocation and operators have a chance to change their bid in the light of this. This procedure does not seem very different from the iterative procedure already followed in timetabling but it has not been adopted in practice to date.

An alternative, explored in this paper, is for the track charging authority to attempt to calculate directly the costs involved. For instance, if a train has to be run at a different time from that desired, it is possible to use studies of the value people place on departure time shifts to estimate the value to its customers of the cost involved. Similarly, the costs of slower speeds may be estimated from passengers' values of time. The Strategic Rail Authority adopted exactly this approach in its Route Utilisation Strategy Studies (now taken over by Network Rail)

A third approach, recommended by Hylen (1998), is to identify sections of infrastructure where capacity is constrained and to charge the long run average incremental cost of expanding capacity. However, this is a very difficult concept to measure (the cost of expanding capacity varies enormously according to the exact proposal considered and it is not easy to relate this to the number of paths created since they depend on the precise number and order of trains run). It may be argued, however, that more appropriate incentives are given to infrastructure managers if they are allowed to charge the costs of investment they actually undertake, rather than for the scarcity resulting from a lack of investment, at least if they are commercially oriented. For short run, marginal cost pricing encourages them to restrict capacity in order to keep price high; whereas a system where a capacity charge reflected actual expenditure on expanding capacity would overcome this problem. EC Directive 2001/14 which governs rail infrastructure charges, seeks to get round this by requiring infrastructure managers to undertake studies to determine the cost of expanding capacity and to test whether this is justified on cost-benefit grounds where scarcity charges are levied.

Given the current degree of excess demand for slots, it is likely that the failure to charge for scarce capacity, together with the under charging for congestion and the exclusion of certain other elements of marginal social cost, is leading to a situation where slots are substantially under priced, compounding the problem of capacity shortage by leading to incentives to train operating companies to run too many trains, to have too strong a preference for frequent short trains rather than less frequent long, and to seek particular routes or timings that are wasteful of capacity.

One counter argument should be considered first, however. That is the argument that the infrastructure of other modes, including road and air, is also not charged for in a way that adequately reflects marginal social cost and in particular, congestion and scarcity costs. Indeed, the study referred to above (Sansom *et al* op cit) found substantial undercharging for the road mode on average, and too little differentiation in the current charges (primarily fuel tax) between locations and times where congestion is a problem and those where it is not. It may therefore be argued that to charge rail operators for these costs when road operators do not have to pay them is counter productive.

This is, however, a very simplistic view of the appropriate approach to such "second best" conditions. The extent to which particular services divert passenger or freight traffic from congested or environmentally damaging roads or airports will differ with the type of service and how heavily it is loaded. The appropriate way of dealing with such second best considerations is therefore to pay train operating companies grants to reflect the benefits elsewhere of diverting traffic from other modes. The government does indeed already pay grants to freight customers designed to attract traffic to rail in such circumstances, related to value of traffic attracted. However, grants for passenger services do not at present vary with or reflect the benefits of the attraction of additional passengers to rail and more could be done to improve incentives here.
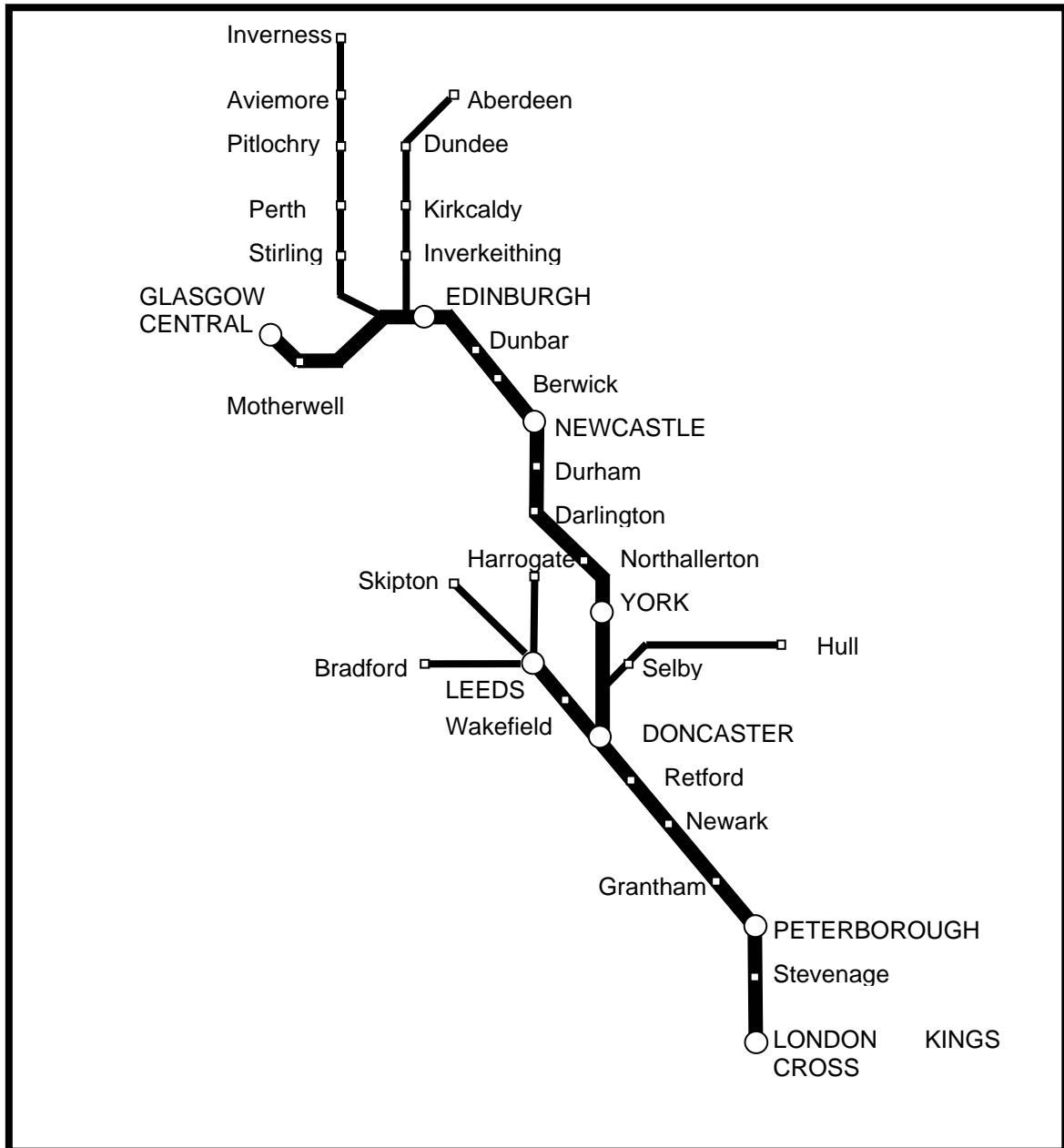
2   The East Coast Main Line

This case study concerns the stretch of the East Coast Main Line from London to Doncaster. The East Coast Main Line forms the principal trunk route from London to Leeds, York, Newcastle and Edinburgh; many trains continue to Glasgow. It is heavily used, particularly between London and Doncaster, which is where the main lines to Leeds, Hull and an important route to Scunthorpe and Grimsby branch off.

The main operator of long distance passenger services on this route and current franchise holder is currently Great North Eastern Railways (GNER), henceforth referred to as "the franchisee". However, a few years ago a new open access operator, Hull Trains, a subsidiary of GB Rail, henceforth referred to as "the open-access operator", was granted access rights to operate through trains between London and Hull (the franchisee only operated the one through train per day required by the franchise agreement; the rest of the time London-Hull passengers had to change trains at Doncaster). The franchisee currently operates a mix of eight car electric and diesel trains capable of 125 MPH, which is the maximum line speed. The open access operator originally operated three car diesel multiple units, with a maximum speed of 100MPH, sometimes operated in pairs to give a maximum of six coach trains, although it is now introducing longer multiple unit trains capable of 125 MPH.

The line from London to Doncaster also carries freight traffic, including between Peterborough and Doncaster container trains operated by Freightliner and by GB Railfreight linking Northern England with the port of Felixstowe and traveling at a maximum of 75 MPH.

There is a shortage of capacity over the Peterborough-Doncaster stretch of the route, which is mainly double track with occasional passing loops. The franchisee would like to increase frequencies and in particular to operate a half hourly service all day between London and Leeds. Its case rests on the fact that it operates long heavily loaded trains and that demand on this, the biggest market on the line, is buoyant. At the time of the study, the open access operator already operated four trains per day in each direction between Hull and London and wished to expand to a two hourly service (it now runs six trains per day each way). Its case is that it offers the convenience of a through service to a city off the main line. At the same time, freight traffic and in particular container traffic to Felixstowe, is growing and there is a demand for more freight paths. Moreover, a further open access operator, Grand Central, applied for paths to run through services from London to Sunderland and Halifax, and has been granted three daily paths to and from Sunderland.

The opportunity cost of paths over this section of track will be the value of the paths in the highest value use. This will obviously vary by time of day. So, to develop a tariff for scarce capacity, we should really examine a whole range of uses of paths for a variety of times of day. However, given the limited time available for this case study, we have examined a much more limited range of options. We have examined the value of an existing peak path each way in each direction to the franchisee and then replaced this with an extra path for the open access operator to compare the values of different uses of this slot. We have also looked very roughly at the value of additional container train paths between Leeds and Felixstowe.

Figure 1: East Coast Mainline route

3   Methodology

3.1   Charging for capacity

The basis of the approach taken here is that operators should be charged for the capacity they use in accordance with the social opportunity cost of that capacity. In order to

implement this approach, it is necessary first to measure the amount of capacity used by each train run and then to estimate its opportunity cost.

For the franchised passenger operator, the impact of changing track access charges is neutralized by the fact that, under the franchise agreement, it is simply passed through to the government as a change in subsidy or premium paid. Where capacity charges may play an important part is in reflecting the opportunity cost of the passenger franchise not using these paths. Currently, other operators only pay the variable part of the infrastructure charge and have no incentive to economize in their use of capacity, for instance by changing speeds, time of day or route.

Thus, the approach we investigate in this paper is construction of a tariff based on the opportunity cost of the slot to the franchisee. If the open access or freight operator requires capacity that would deprive the franchisee of more than one slot (for instance, because their trains are slower than those of the franchisee), then they would be charged for the appropriate number of slots. Since the franchisee is known and is required to make data available to the regulator, this approach to charging should be feasible. Of course, if there are several other operators competing for the slot and they all have higher values than the franchisee, then this will understate the true opportunity cost of the slot. However, basing charges on the identity of unknown possible new entrants appears difficult, at least until they start operating and data becomes available.

The opportunity cost of a slot for this type of service can be estimated as the sum of:

- the additional amount of traffic attracted to rail by the presence of this train multiplied by the price it pays

- the consumers' surplus to rail users as a result of the additional quality and capacity provided by the train (including reduced crowding on other services)

- the savings of external costs to road users and the public at large from the train attracting passengers from road.

- less the train operating, infrastructure and external cost savings from failing to run this train.

Given that this route is one where a major conflict appears to exist between plans for passenger services and the desire to expand freight services, it is interesting to do some exploratory estimates for the net benefit of an additional freight service. Table 1 suggests that non bulk freight services in 1998 were typically earning revenue some £1.65 in excess of marginal social cost. For bulk commodities, a mean load of the order of 1400 tonnes is now normal. For a container train, the load may only be around 450 tonnes. For bulks therefore, around 45 hgvs will be removed from the road system; for container trains, (which we have assumed for this case study), it may be more like 30 hgvs. The revenue figures presented below are averaged over loaded and empty movements, and the other benefits accrue whether loaded or empty hgvs are removed from the roads. Although on average, the probability of a road vehicle returning empty is much lower than for rail, for these specific traffics, it is postulated that the difference is not very great. Therefore, the benefits shown in the table will roughly apply whether the train is loaded or empty.

There is a benefit of 23.5p per vehicle km from relieving the roads of this traffic. Given our assumption above that there are around 30 hgvs per container train, there is a benefit of £7 (ie 23.5p*30) per container train to add to the £1.65 excess of rail revenue over costs, as

shown in table 1. So the total benefit of the freight service is £8.65 per train km. However, because of their slower speed, it is likely that they require more than one standard slot per train run. Also, it is more likely that freight services may be retimed or rerouted, as opposed to the traffic being totally lost, in which case the relevant net benefit from an ECML path may be much lower than that quoted above. Moreover, there is less certainty that a slot allocated to freight traffic will actually be used given short term fluctuations in traffic. If the slot is only actually used on 50% of occasions, then its value is obviously halved.

	Category	
Social Costs	Rail Freight (Container) trains	Road Freight Vehicles
Marginal infrastructure usage	0.88	0.0869
Air pollution	0.684	0.0452
Noise	0.367	0.0222
Climate change	0.328	0.0179
Congestion		0.2452
Accident costs		0.0120
Private Costs		
Train operating cost	9.70	
Total	**11.96**	**0.4293**
Revenue	13.61	
Tax Revenue		0.195
–Surplus of revenue over cost	+1.652	-0.235

Table 1: Marginal cost and revenue analysis for freight 1998 £/ train/ vehicle km

Note: Road figures are taken from the mid points of values for marginal cost and revenues for articulated lorries given in tables 7.4 and 7.5 of the Sansom report. Rail figures are taken from 8.7 and 8.8 of the report. Road freight infrastructure charges are annual license duty and fuel tax. The operating costs and revenues of the road hauler are assumed to exactly offset each other, and are therefore excluded.
Source: Sansom *et al* (2001).

3.2   The PRAISE Model

3.2.1 Overview

The value of the passenger slots will be estimated using the PRAISE model. The PRAISE (Privatized Rail Services) model was developed at the Institute for Transport Studies, University of Leeds to look at the potential for open access competition following the privatization of rail services (Whelan *et al*, 1997; Preston *et al*, 1999). More recently, the model has been re-written and developed to be capable of assessing demand and costs for

small networks of stations incorporating the services of up to five operators, each with 10 different ticket types (Whelan, 2002). The software comprises a demand model, a cost model and an evaluation model.

PRAISE forecasts demand for individual services and ticket type, taking account of fares, journey times, desired departure times and overcrowding, so it is very useful for looking at issues concerning capacity, detailed timetabling and fares and ticket restrictions, as well as competition between different operators. In this case, it will forecast the extent to which changes in the timetable will lead to changes in rail passenger traffic, taking account of the precise times of the trains affected, the possibility of passengers taking other trains in the timetable or ceasing to use rail at all, and the changes in the fares and levels of crowding passengers face on the different options.

There are three stages to the calibration of the demand model. The first involves the estimation of the generalized cost of travel for each return service and ticket combination. The second involves setting the "scales" of the choice model so that it replicates known elasticities of demand. The third involves calibrating ticket specific constants to ensure that the base market shares can be replicated.

The cost model employs a cost accounting approach incorporating costs that are related to operating hours, costs that are related to train kilometers and fixed costs. Costs can be varied by operator and rolling stock type based on figures from the Rail Industry Monitor (TAS, 2004) and can be combined with estimates of revenue to generate forecasts of operator profitability.

The model generates output that can be used in a formal appraisal system. This output includes passenger demand, passenger distance, operator revenue, operator costs, profitability, user benefits (consumer surplus), overcrowding, and diversion to and from other modes in terms of passenger numbers and passenger distance. The demand, cost and appraisal models are outlined in more detail in the following sections.

We run the following scenarios in PRAISE, from which we can get some idea of how the value of a slot to the franchisee varies with time of day:

Removal of existing 9:05 Peak Franchisee Service. Given the existing pricing policy, in which this is the first train on which the off-peak "saver" fare applies, this is the busiest Leeds bound Franchisee service, and as such can be used as a benchmark against which we can judge the impact of adding other services.

- Removal of existing 8:05 peak franchisee service.

- Introduction of additional of 8:35 peak franchisee service.

- Introduction of additional of 9:35 peak franchisee service.

- Introduction of additional off-peak 15:40 service.

The additional slots are those needed to provide a half hourly service on this route, as desired by the franchisee. In order to get some idea of the value of slots to other operators we also test:

- Introduction of additional peak and off-peak open access services.

- Introduction of additional container services, for which we apply the values estimated in table 1 above.

3.2.2 Lower level of the demand model

The demand model has a hierarchical structure and works at the level of the individual traveler. The lower level of the model uses a nested logit to assign a probability that a given traveler will choose a particular ticket, and outward and return service combination. The nesting of the journey options into three groups for full, season and reduced tickets are to accommodate the different elasticities between these ticket types. By aggregating the ticket and service probabilities over a representative set of simulated passengers, the model is able to forecast market shares for each service and ticket combination. To allow for the fact that changing fares and services will change the overall demand for rail, the upper level of the model is structured to allow the rail market to expand or contract according to changes in average fares and journey times multiplied by the appropriate elasticities. For each origin-destination pair on the network in a given operational period (for example, a typical weekday), the model generates a sample of i simulated individuals with given tastes expressed in terms of relative journey attribute valuations (for example, the value of time and adjustment time) and elasticities of demand. Individual level heterogeneity is generated through differences in desired outward and return leg departure times. These are selected for an individual from a distribution which captures the characteristics of the journey type and purpose, for example, short distance commuters are more likely to have desired departure times early in the morning and departure times around the evening peak. When making a rail journey, a passenger will often have a choice from a number of different services and ticket types. For example, they could choose to travel in the peak using an unrestricted ticket or in the off-peak using a less expensive restricted ticket.

If we know when the passenger would ideally like to travel, we can estimate a generalized cost for each option n (service and ticket combination) available in the following way

$$GC_n = F_n + (vot * GJT_n) + CP \tag{1}$$

where:

- F is the return fare

- GJT is the generalized journey time (minutes)

- vot is the behavioral value of time (pence per minute)

- CP is a crowding penalty (pence)

The generalized journey time is calculated as:

$$GJT_n = IVT_n + \left(\frac{vat}{vot} * AT_n \right) + IP_n + \left(\frac{vwt}{vot} * OVT_n \right) \tag{2}$$

where:

- IVT is in-vehicle time (minutes)

- *AT* is schedule adjustment time (minutes). This is the difference between a passenger's most desired time of departure and the actual timetabled departure time.

- *vat* is the behavioral value of schedule adjustment time (pence per minute)

- *IP* is an interchange penalty (minutes)

- *OVT* is out of vehicle time (minutes), that is, the time spent waiting between interchanges

- *vwt* is the behavioral value of time spent waiting

The attributes included within the generalized cost expression are for the most part well known and there is a wealth of literature providing evidence on relative attribute values (see for example, the Passenger Demand Forecasting Handbook (PDFH, ATOC, 2002).

The generalized cost expression we use relates to a return journey and therefore contains generalized cost elements for both the outward and return legs. This feature is important when there is more than one operator and tickets are operator specific.

Passengers on crowded or overcrowded trains will typically experience discomfort associated with having to stand or sit in cramped conditions. The level of discomfort varies according to whether the passenger is sitting or standing, the degree of overcrowding, the length of the journey and the type of journey being made (for example, commuters may be used to overcrowding on short journeys). Much of the published research on passengers' valuation of overcrowding uses stated preference techniques to assess the trade-off between fares, times and crowding to derive monetary or time estimates of overcrowding penalties (see ATOC, 2002). Penalties vary depending on the type of traveler; commuters, who may be used to overcrowding on short journeys, have the lowest penalties. Business travelers who may regard it as degrading, suffer the highest penalties.

When making a rail journey, a passenger will often have a choice from a number of different services and ticket types. For example, they could choose to travel in the peak using an unrestricted ticket or in the off-peak using a less expensive restricted ticket.

If we know when the passenger would ideally like to travel, we can estimate a generalized cost for each option *n* (that is, service and ticket combination) available and assign each a probability that it will be chosen:

$$P_n = P'_r \, P_{n|r} \tag{3}$$

where P'_r is the probability of choosing rail and $P_{n|r}$ the probability of choosing option *n* conditional on the choice of rail.

The probability that an individual will choose a given service and ticket combination *n* conditional that they chose rail, *r*, is:

$$P_{n|r} = \frac{\exp(U_n)}{\sum_{n' \in N} \exp(U_{n'})} \tag{4}$$

where U_n is the utility of option n, which is given by:

$$U_n = -\lambda_t GC_n + ASC_t \qquad (5)$$

where

- $U_{n'}$ is the utility of option n'

- GC is the Generalized Cost for option n

- ASC_t is an Alternative Specific Constant for ticket t (from ticket/service combination n) to ensure that the model is able to correctly retrieve ticket market shares in the base period.

The alternative specific constants for each ticket type are initially set to zero and the model is then run to generate forecasts of the share of each ticket type. These forecasts are then compared with actual ticket sales data and a new set of model constants derived. These constants capture the impact of the range of factors that influence passengers' choice of tickets that are not already included within the generalized cost expression. They are derived as follows:

$$ASC_t^{new} = ASC_t^{base} - \lambda_r \ln\left(\frac{MS'_t}{MS_t}\right)$$

$$ASC_t^{new} = ASC_t^{base} - \lambda_r \ln\left(\frac{P'_t}{P_t}\right) \qquad (6)$$

where:

- ASC_t^{base} is the base constant for ticket type t

- ASC_t^{new} is the new constant for ticket type t

- MS'_t is the forecast market share of ticket type t

- MS_t is the actual market share of ticket type t.

The scale parameter associated with each ticket type, λ_t governs the sensitivity of choice to changes in the generalized cost. As the value of λ_t approaches zero, all N options have an equal chance of being chosen whereas as the value of λ_t increases, the probability that the option with the lowest generalized cost tends to one. These scales can be set to replicate known fare and GJT elasticities of demand for full, season and reduced tickets.

The choice modeling hierarchy is repeated for a sample of individuals drawn from known desired departure time profiles. The market share for each option (service and ticket type) is taken as the average probability for each option over all individuals in the sample.

3.2.3 Upper level of the model

The upper level of the model is concerned with the overall size of the rail market. This is modeled using an elasticity approach based on changes in average fare and GJT values in the following way

$$P'_r = P_r * \left(\frac{F}{F'}\right)^{\eta_f} * \left(\frac{GJT}{GJT'}\right)^{\eta_{gjt}} \tag{7}$$

where:

- P'_r is the new probability of choosing rail

- P_r is the base probability of choosing rail

- F is the base average fare level

- F' is the new average fare level

- GJT is the base average generalized journey time

- GJT' is the new average generalized journey time

- η^f is the appropriate fare elasticity taken from the PDFH (ATOC, 2002)

- η^{GJT} is the appropriate GJT elasticity taken from the PDFH.

In accordance with PDFH, GJT elasticities vary by distance, average speed and the degree of competition from other modes, whilst fare elasticities vary by type of flow (that is, to or from London) and distance.

3.2.4 Calibration and forecasting

The final stage of the calibration of the demand model includes a number of iterations to account for changes in distribution of passengers on trains through the overcrowding penalty. Given the degree of crowding on trains during the peak on the network, this is an important development in the modeling. Following the estimation of the generalized cost of each option, the model is applied to generate forecasts for each operator and ticket type.

To allow for the fact that changing fares and services will change the overall demand for rail, the upper level of the model (equation 7) is structured to allow rail volumes to expand or contract according to the overall quality of rail, as defined by changes in generalized journey times and fares.

Following the estimation of the new generalized cost of each option, the model is applied to generate new forecasts for each operator and ticket type. The process is iterated again to incorporate effects of crowding penalties.

3.2.5 External costs and benefits

Whilst PRAISE will yield consumer surplus, operating profits, modal switch values and changes in vehicle kilometers, it does not carry out a full appraisal incorporating valuations

of external costs. These external costs comprise those imposed by rail itself, for which we use the estimates presented in Table 1 above and those imposed by other modes of transport whose volumes are changes by the change in rail frequencies.

For external costs and benefits of other modes of transport, we again use values from the study of Sansom *et al* (2001). It should be noted that strictly these values were estimated for 1998, whereas our other data applies to 2002. To apply these values, we need to know how much traffic transfers to or from road and the types of road and time of day in question.

The change in rail passenger trips can be used to calculate the modal shift between rail, car, coach and not travel or new journeys. An integral part of these calculations are the application of diversion factors to the change in passenger trips. For example, if the number of rail trips is assumed to have increased by 10,000 per year, diversion factors can be used to ascertain where those journeys have come from. In the appraisal, the following diversion factors (Table 2) were used to estimate the sources of new rail journeys and vice versa.

Diversion Factors	%	Sources of New Rail Journeys	
Car	68%	Car	6,800
Coach	24%	Coach	2,400
New	8%	New	800

Table 2: Diversion factors & sources of new rail journeys

Source: Train Operating Company Figures (1998).

From the table, it is clear that 6,800 of the 10,000 new rail journeys are being made by people who used to travel by car; that 2,400 trips are made by former bus/coach passengers; and that 800 journeys are new in that they were not made previous to the introduction of the new timetable. This information can be taken forward and used to calculate a number of the impacts outlined in the appraisal framework.

To calculate the modal shift in terms of car and coach vehicle kilometers (kms) requires the average loadings of both car and coach vehicles to be taken into account, alongside the length of the trips made by both modes. In the case of car, a loading factor of 1.6 (taken from the Transport Economics Note, DfT, 2003) has been used and in the absence of any supporting data, we have assumed for coach a loading factor of 25. This allows the number of car and coach journeys to be calculated (see Table 3) following on from the figures presented in Table 2.

	Modal Switch (passenger journeys)	Modal Switch (vehicle journeys)
Car	6,800	4,250
Coach	2,400	198

Table 3: Calculating modal switch in terms of journeys

To calculate the total number of car and coach vehicle kms that have been switched, the total distance of the trip needs to be factored in. These figures then need to be multiplied by the number of journeys for each mode to calculate the total modal switch in terms of vehicle kms.

This information can be taken forward and used to calculate the external cost changes on other modes. All the factors used for the calculation of the environment have been taken directly from a report carried out by ITS for the DETR which examined surface transport costs and charges for Great Britain for 1998 (Sansom *et al*, 2001). The report calculated the costs per vehicle kilometer for road and rail travel. In cases where high and low marginal cost values have been reported, we have taken the mid point between these two values. The UK average values for environmental factors are presented in Table 4. These values have been uplifted to 2003 prices using the RPI.

Impact Type	Coach	Car	Passenger Rail	Freight Rail
Noise	0.021	0.0027	0.122	0.170
LAQ	0.093	0.0053	0.279	0.166
Greenhouse Gases	0.014	0.0030	0.067	0.131
Safety	0.052	0.011		

Table 4: UK average values of environmental factors (£s in 1998 prices and values)

The table below (5) shows the figures below for infrastructure costs, from Sansom *et al* (2001). Road Infrastructure costs are based on the average values for vehicle kms, vehicle type and road types.

Impact Type	Road (Car)	Road (Coach)	Rail (Intercity)	Rail Freight
Infrastructure costs	0.0006	0.060	1.116	1.19

Table 5: UK average values of infrastructure costs (£s per vehicle or train km in 1998 prices and values)

We assume that infrastructure costs have already been charged to the train operator through the variable access charge.

For car and coach travelers, the change in user benefits is reflected by the change in congestion costs that they incur. The costs of congestion are outlined in the table below (6).

Impact Type	Coach	Car
Congestion	0.1522	0.0898

Table 6: UK average values of external congestion cost (£s per vehicle km in 1998 prices and values)

In the absence of sufficient revenue data, we took fares information from the National Express website, which allowed us to approximate a marginal profit figure of 1.6 pence per

passenger kilometer. This has been based on an assumed journey length of 150km and that half of total costs are fixed. We assumed coach service costs would expand, maintaining existing load factors. This will increase external costs but there will be benefits to existing users from increased frequencies (the Mohring effect), taken from Sansom *et al*, of 14.7 pence per vehicle km.

The impact of indirect tax directly affects government revenues. For cars, the government levies fuel duty and VAT on fuel duty. Rail and coach travel are not subject to VAT, so VAT not paid on fares expenditure which would have otherwise incurred VAT has to be calculated as a cost of these modes. Values per average UK vehicle kms have been taken from the Sansom *et al* (2001) publication and are presented in Table 7.

Mode	Car	Coach
Fuel Duty	0.0386	0.0526
VAT on Fuel Duty	0.0068	0.0092
VAT not paid	Not applicable	0.1344

Table 7: UK average values for indirect taxes per vehicle kms (£s-1998)

VAT is charged at 17.5%, so changes in VAT not paid from Rail can be derived from the change in rail revenues.

With regards to subsidy, we have assumed there will not be any effect on subsidy payments. The coach services in question in this case study are unsubsidized express coach services, whilst the rail services are either commercial expansions of existing franchise agreements or open access commercial operations .

4   Data description

4.1   Stations for case study

We are primarily interested in flows between Doncaster and London but obviously need to consider other flows which are based on, or at some point using the East Coast Mainline to consider capacity issues, overcrowding and competition from other operators.

In order to represent demand that has its origin or destination outside of the immediate study network, we define a number of "buffer" stations. These allow us to model peripheral demand without having to model the rail network as a whole, and as such represent an aggregation of demand from stations that feed into the network.

When using a buffer station, we aggregate the demand from all the stations in the buffer zones but use the fares, distances, elasticities, values of time and market shares for the buffer station. For this reason, the named buffer station is chosen to be representative of the average distance between stations in the buffer zone and other stations.

Results will be presented based on all combinations of flows between the following proposed set of stations.

- *Leeds*; this acts as a buffer station for flows onto the ECML network North West of Doncaster, such as Leeds, Bradford, Wakefield and Huddersfield, Ilkley and Skipton

- *Newcastle*; A buffer station for all points north of Doncaster on and using the ECML, such as Scarborough, York, Harrogate, Thirsk, Northallerton, Darlington, Durham, Berwick, Edinburgh etc.

- *Hull*; with demand aggregated with that of stations North East of Doncaster, including Hull, Beverley, Bridlington

- *Doncaster*- incorporating demand from Selby, Goole, Scunthorpe Grimsby and Cleethorpes

- *Peterborough* ; incorporating Northbound flows from East Anglia which use the ECML such as Cambridge, Hitchin, Stevenage and Stansted Airport

- *Grantham*; A buffer station including Newark, Retford, Skegness, Boston and Sleaford

- *London Kings Cross*

- *South East Buffer*; including demand to 100 MOIRA stations in the South East

We are constrained in the number of stations we choose by the ever increasing run times associated with increasing numbers of outward-destination pairs.

4.2   Rail industry data

Demand figures, fares and market shares are based on typical inter city demand figures derived from 2003/4 ticket sales data.

Average costs per train km are based on typical inter city costs from the *Rail Industry Monitor* (TAS, 2004) and figures from Network Rail. As we are focusing on changes in demand, we do not need to explicitly calculate the fixed element but do make sure that the variable cost covers the subset of total costs that would vary by distance. In order to derive this, we have considered the following as short run variable costs:

- Variable Track access charges

- Electricity charges

It is assumed that staffing and leasing costs and other costs would remain the same.

Whether rolling stock leasing costs should be regarded as fixed is open to question. In the short run, TOCs regard these as fixed because they have rolling stock leasing contracts lasting a number of years. But obviously if a TOC ceased operating on this route it would seek to redeploy the units released elsewhere. Also, whilst the individual slots considered here may not affect rolling stock requirements, if a TOC expanded its service more generally it may well need to expand its fleet.

Generalized Journey Time Elasticities were taken from the industry standard Passenger Demand Forecasting Handbook (PDFH), based on Revealed and Stated Preference research.

Values of time and adjustment time were also taken from the PDFH, based on Revealed and Stated Preference research. The value of time used was a weighted average of business, leisure and commute time values.

Crowding Penalty valuations are taken from the PDFH and are in pence per crowded minute of journey, varying by route and degree of crowding.

5   Results

Because of the confidentiality of some of the data, it is not possible to reproduce detailed quantitative results here, so all monetary values are given as relative, in percentage terms, to the full economic value of the existing use of the London peak (9:05 from Leeds) slot by the franchisee, (necessarily itself 100%) as presented in table 8. For example, the loss in operators profits in table 8 accounts for 34% of the overall welfare loss. The 9.05 is the most heavily loaded train because it is the first train on which the cheaper "Saver" fare is valid. It can be said that its scarcity charge would be very substantial relative to existing track access charges.

	Rail	Car	Coach
Change in trip kms	*-296.5*	*200.2*	*67.0*
Change in vehicle kms	*-1.0*	*125.1*	*2.7*
Environmental impacts			
1.1 Noise	0.5	-1.5	-0.5
1.2 LAQ	0.0	-3.0	-2.3
1.3 Green house	0.3	-1.7	-0.3
1.4 Safety	0.0	-6.1	-1.3
Infrastructure costs		-0.3	-1.5
Tax revenues	9.3	21.8	-2.2
User benefits			
Consumer surplus	-26.3		
Congestion		-54.8	-4.2
Mohring effect			3.7
Operator's profits	-34.0		4.4
Total welfare change	**-100.0**		
Full economic value of slot	**100.0**		

Table 8: Effect of removal of peak franchisee's 9:05 service

Note: All monetary values as percentage of net social benefit of use of peak slot by franchisee.

All distances are per change in train km, for example in table 8 there is a reduction of 297 rail trip kms for each train km removed. Although we cannot present an explicit value for scarcity cost, we can use relativities presented to show how the different uses of any extra paths compare to the use of a peak path by the franchisee.

We look at seven passenger scenarios. Scenario 1 is the removal of the franchisee's 9:05 service as shown in table 8. Scenario 2 is the removal of the franchisee's 8:05 peak

service. It is appreciated that in practice the loss of a slot is likely to lead to withdrawal of a service in the opposite direction as well. Scenario 2 is the removal of the earlier 8:05 peak service from Leeds. Scenarios 3 and 4 involve an extra peak train at 8:35 and 9:35 respectively for the franchisee. Scenario 5 is an extra off-peak train leaving Leeds at 15:40, halfway between two hourly services. Scenarios 6 and 7 are the introduction of an extra open access service in a peak and off-peak slot respectively. Detailed results for scenarios 2 to 7 are given in the appendix.

Table 9 summarizes the results of the seven passenger train scenarios and also presents the value of a freight train's use of a peak slot as scenario 8. The difference between the values in the first two columns represents the difference between the impact on the profitability of the operator of the slot and that on all operators allowing for diversion of passengers between trains.

Scenario	Value of operators' Profits (as % of Economic value of 9:05 Peak Slot)	Value of profits to operator of slot (as % of Economic value of 9:05 Peak slot)	Net social benefits (as % of Value of 9:05 Peak Slot)
1: Removal of franchisee's peak train (9:05)	34.0	55.2	100.0
2: Removal of franchisee's peak train (8:05)	4.4	18.6	31.4
3: Extra franchisee peak train (8:35)	7.3	15.9	27.1
4: Extra franchisee peak train (9:35)	8.9	19.2	31.7
5: Extra franchisee off peak train (15:40)	17.1	22.2	29.5
6: Extra open-access peak train (7:19)	8.3	43.5	22.2
7: Extra open-access off-peak train (11:06)	0.6	14.5	5.6
8: Extra Freight Service	11.5		60.0

Table 9: Summary of private and social slot values

Note: As percentage of net social benefits of use of 9:05 peak slot by franchisee.

In terms of net social benefits, scenario 1, the franchisee's current use of the 9:05 peak path gave the highest values for passenger use. The 8:05 peak path in scenario 2 appears to be less valuable than the 9:05, representing only 31% of the value. Extra peak services either side of the 9:05 in scenarios 3 and 4 have a significant impact in terms of profits and net social benefits. The 8:35 and 9:35 services would yield 27% and 32% of the value of the 9:05 service, respectively. The off-peak slot in scenario 5 has a similar effect. The use

of the peak slot by the open access operator in scenario 6 has a smaller net benefit than additional franchisee services and an extra off-peak slot for the open access operator, as shown in scenario 7, is even less valuable.

In terms of private profitability to the operators running the slot, the highest values are 55% (of the total net social benefit of the 9:05 peak slot) to GNER for the 9:05 peak slot and 44% to the open access operator for the peak 7:19 service. The lowest is 15% for the off-peak open access service. In the case of the open access operator's use of the slot, the increase in the operator's profits is mainly at the expense of the franchisee. There were very large differences between private and social profitability, suggesting that auctioning without the payment of explicit subsidies to operators to reflect the social benefits of the use of the paths would not necessarily give the optimum result. Our figures do not consider the value of the slots in the context of the provision of an overall regular interval timetable. Studies have shown (see Johnson, Nash and Tyler (2006)) regular interval timetables to confer extra benefits to passengers over and above those from changes in journey times due to more convenient journey planning, so we may have underestimated the value of additional off-peak slots to the franchisee where these make regular interval services possible.

Freight valuations do not distinguish between peak and off peak, as derived freight demand is not determined from time of day issues.

Although we cannot report the freight results without indirectly revealing the passenger service values, the imputed benefit from using the paths for freight was 60% of the existing use of the peak 9:05 slot for passenger use. It is not clear exactly how much capacity would be required for an additional freight train. Whilst it may be assumed that a freight train typically requires twice the capacity of a passenger train, many container trains are routed off the ECML between Newark and Doncaster, and there are only relatively short sections where there are only two tracks and some of these have passing loops. Since these also run at relatively high speeds, the opportunity cost of container trains may not be very high in terms of number of paths lost. However, for more slowly moving bulk trains, which run to demand, the opportunity cost of reserving capacity may be prohibitive.

6   Conclusions

This paper considers the issue of charging for scarcity by calculating the value of scarcity in terms of the opportunity cost of making paths available to the main franchised operator on an intercity route in Britain. The institutional arrangements behind franchising mean that the Regulator has data for the franchisee from which he can calculate the opportunity cost to the franchisee of the use of a slot and thus the scarcity value. Our approach is that if any other operators wish to use capacity desired by the franchisee, they should pay the opportunity cost of the use of the path by the franchisee. Clearly, in trying to evaluate what would be the outcome of the imposition of a scarcity charge, we have had to use information on potential entrants but the ECML is an example of a line with open access competition so we have information on competing passenger and freight operators.

We have used typical industry data applied with a detailed rail passenger simulation model which produces estimates of revenue, costs, consumer surplus and diversion to/from other modes. Estimates of changes in external costs are then made to derive results for the overall social benefits of alternative allocations of capacity.

Our approach is, however, limited in its scope, deriving scarcity values based on the value of slots to the current dominant operator in the face of small changes to an existing timetable and where there is good knowledge of the demand and cost characteristics of their operations. This corresponds to the typical situation were there is a dominant franchisee and potential new entrants but would be less helpful if there were extensive on track competition between passenger and/or freight operators.

Because of the nature of the franchise agreement in Britain, this charge would only affect open access passenger and freight operators, reflecting the impact of their decision to operate on the services of the passenger franchise. If it was desired to apply the charges also to the franchisee if they sought additional slots, then it would be necessary to set the charge somewhat below the value to the franchisee to avoid the risk that the slots may be wasted. The results seem to confirm the view that existing variable charges for the use of infrastructure on key main lines where capacity is scarce are much too low as a result of the neglect of scarcity in the charges set.

In terms of passenger trains, the existing use of the 9:05 peak slot by the inter-city operator has the highest social value. We judged the impact of other services against this. The highest value extra service was for the peak services offered by the existing franchisee. It would appear from these calculations that whilst there is no case for taking existing slots from the incumbent, scarcity charges based on net social benefits of allocating additional slots may be significantly lower than the value of an existing slot, and not vary greatly between peak and off peak. However, it did not appear that open access operators would in general be able to compete with the franchisee for slots.

However, such charges would only have the appropriate impact if subsidies representing the net social benefits of attracting additional traffic to rail were paid in a way which directly added to the benefits of running additional train kilometers to the operator and operators were only rewarded for attracting additional revenue to rail rather than diverting revenue from existing trains.

Container freight trains may also gain paths provided they move at reasonable speeds and run on most occasions they are allocated a path. Bulk freight, which would require at least two passenger paths and which typically only uses the path on around 50% of occasions for which it is reserved, would be unlikely to win paths on the ECML.

We believe that this paper suggests that the imposition of scarcity charges based on the value of slots to the franchisee is both feasible and likely to be socially beneficial. However, more work is needed on exactly what the tariff should look like and what its overall effects would be.

7   References

ATOC (2002) "Passenger Demand Forecasting Handbook," Association of Train Operating Companies, London.

Bowker, R (2002) "Britain's Railway – Time for a New Radicalism. (The Sir Robert Reid Railway Lecture 2002)," The Institute of Logistics and Transport, London.

Department for Transport (2003) "Transport Economics Note," available at http://www.webtag.org.uk/webdocuments/3_Expert/5_Economy_Objective/3.5.6.htm

Hylen, B. (1998) "An Examination of Rail Infrastructure Charges," Final Report for the European Commission, DG VII. Prepared by NERA, Nomisma, VTI, IVE ENPC & BPM, London.

Johnson, Nash and Tyler (2006) "Forecasting and Appraising the Impact of a Regular Interval Timetable," *Transport Policy*, 13: 349-366.

Nilsson, J-E (2002) "Towards a Welfare Enhancing Process to Manage Railway Infrastructure Access," *Transportation Research A*, 36: 419-436.

ORR (2000) "Periodic Review of Railtrack's Access Charges: Final Conclusions," Office of the Rail Regulator, London.

Preston, J., Whelan, G.A., and Wardman, M.R. (1999) "An Analysis of the Potential for On-track Competition in the British Passenger Rail Industry," *Journal of Transport Economics and Policy*, 36: 77-94.

Sansom, T., Nash, C., Mackie, P., Shires, J.D. and Watkiss, P. (2001) "Surface Transport Costs and Charges: Great Britain 1998," A Report for the DETR.

TAS (2004) "Rail Industry Monitor," available at http://www.tas-passtrans.co.uk/Rail+Industry+Monitor

Whelan, G.A., Preston J.M., Wardman, M., and Nash, C.A. (1997) "The Privatisation of Passenger Rail Services in Britain: an Assessment of the Impacts of On-the-Track Competition," Presented to European Transport Forum, PTRC 1997.

Whelan, G.A. (2002) "Analysing Impact of Changes in On-Rail Competition – Model Development," Final Report to the Strategic Rail Authority.

8 Appendix

	Rail	Car	Coach
Change in trip kms	*-119.0*	*80.3*	*26.9*
Change in vehicle kms	*-1.0*	*50.2*	*1.1*
Environmental impacts			
1.1 Noise	0.5	-0.6	-0.2
1.2 LAQ	0.0	-1.2	-0.9
1.3 Green house	0.3	-0.7	-0.1
1.4 Safety	0.0	-2.5	-0.5
Infrastructure costs		-0.1	-0.6
Tax revenues	4.1	8.7	-0.9
User benefits			
Consumer surplus	-12.0		
Congestion		-22.0	-1.7
Mohring effect			1.5
Operator's profits	-4.4		1.7
Total welfare change	**-31.4**		

Table A1: Effect of scenario 2 – Removal of peak franchisee's 8:05 service

Note: All monetary values as percentage of net social benefit of use of peak slot by franchisee.

	Rail	Car	Coach
Change in trip kms	*110.4*	*-74.5*	*-24.9*
Change in vehicle kms	*1.0*	*-46.6*	*-1.0*
Environmental impacts			
1.1 Noise	-0.5	0.6	0.2
1.2 LAQ	0.0	1.1	0.9
1.3 Green house	-0.3	0.6	0.1
1.4 Safety	0.0	2.3	0.5
Infrastructure costs		0.1	0.6
Tax revenues	-4.7	-8.1	0.8
User benefits			
Consumer surplus	6.7		
Congestion		20.4	1.6
Mohring effect			-1.4
Operator's profits	7.3		-1.6
Total welfare change	**27.1**		

Table A2: Effect of scenario 3 – extra peak 8:35 franchisee service

Note: Values as percentage of net social benefit of current use of peak slot by franchisee.

	Rail	Car	Coach
Change in trip kms	*142.9*	*-96.5*	*-32.3*
Change in vehicle kms	*1.0*	*-60.3*	*-1.3*
Environmental impacts			
1.1 Noise	-0.5	0.7	0.3
1.2 LAQ	0.0	1.4	1.1
1.3 Green house	-0.3	0.8	0.2
1.4 Safety	0.0	2.9	0.6
Infrastructure costs		0.2	0.7
Tax revenues	-4.9	-10.5	1.1
User benefits			
Consumer surplus	4.6		
Congestion		26.4	2.0
Mohring effect			-1.8
Operator's profits	8.9		-2.1
Total welfare change	**31.7**		

Table A3: Effect of scenario 4 – extra peak 9:35 franchisee service

Note: Values as percentage of net social benefit of current use of peak slot by franchisee.

	Rail	Car	Coach
Change in trip kms	*92.9*	*-62.7*	*-21.0*
Change in vehicle kms	*1.0*	*-39.2*	*-0.8*
Environmental impacts			
1.1 Noise	-0.5	0.5	0.2
1.2 LAQ	0.0	0.9	0.7
1.3 Green house	-0.3	0.5	0.1
1.4 Safety	0.0	1.9	0.4
Infrastructure costs		0.1	0.5
Tax revenues	-6.4	-6.8	0.7
User benefits			
Consumer surplus	4.0		
Congestion		17.2	1.3
Mohring effect			-1.1
Operator's profits	17.1		-1.4
Total welfare change	**29.5**		

Table A4: Effect of scenario 5 – extra off-peak 15:40 franchisee service

Note: Values as percentage of net social benefit of current use of peak slot by franchisee.

	Rail	Car	Coach
Change in trip kms	93.5	-63.1	-21.1
Change in vehicle kms	1.0	-39.4	-0.8
Environmental impacts			
1.1 Noise	-0.6	0.5	0.2
1.2 LAQ	0.0	1.0	0.8
1.3 Green house	-0.3	0.6	0.1
1.4 Safety	0.0	2.1	0.4
Infrastructure costs		0.1	0.5
Tax revenues	-2.9	-7.5	0.8
User benefits			
Consumer surplus	0.4		
Congestion		19.0	1.4
Mohring effect			-1.3
Operator's profits	8.3		-1.5
Total welfare change	**22.2**		

Table A5: Effect of scenario 6 – extra peak 7:19 open-access service

Note: values as percentage of net social benefit of current use of peak slot by franchisee.

	Rail	Car	Coach
Change in trip kms	32.8	-22.1	-7.4
Change in vehicle kms	1.0	-13.8	-0.3
Environmental impacts			
1.1 Noise	-0.6	0.2	0.1
1.2 LAQ	0.0	0.4	0.3
1.3 Green house	-0.3	0.2	0.0
1.4 Safety	0.0	0.7	0.2
Infrastructure costs		0.0	0.2
Tax revenues	-1.5	-2.6	0.3
User benefits			
Consumer surplus	1.4		
Congestion		6.7	0.5
Mohring effect			-0.4
Operator's profits	0.6		-0.5
Total welfare change	**5.6**		

Table A6: Effect of scenario 7 – extra off-peak 11:06 open-access service

Note: Values as percentage of net social benefit of current use of peak slot by franchisee.