Scoping study for scarcity charges

Final report for the Office of Rail Regulation

Professor Chris Nash and Daniel Johnson

Institute for Transport Studies, University of Leeds

Jonathan Tyler
Passenger Transport Networks, York

December 2005
Revised February and March 2006

1. Background

1.1 The existing structure of rail access charges in Britain, as in most countries, leads to a low marginal charge for the operation of additional trains, despite the introduction of a capacity charge based on congestion costs at the last periodic review. In cases where capacity is scarce, and there are demands on the system which cannot be met, this marginal charge fails to reflect the opportunity cost of forcing other traffic to travel by a less favoured route, at a less favoured time or not to use rail at all. The argument in favour of scarcity charges is that introducing a scarcity charge to cover this opportunity cost would help either in ensuring that slots were allocated more efficiently in the short run, or in terms of the longer term planning of services. Such charges are permitted under European Directive 2001/14, which recognises that they should take the form of reservation fees, paid whether the capacity is actually used or not. It is a condition of such charges that plans to enhance capacity are prepared and proceed unless a cost-benefit analysis shows they are not worthwhile.

1.2 It is important to distinguish between scarcity and congestion. Congestion arises from delays caused by trains running close to each other. In a planned system such as a railway the timetable is designed to prevent this from happening, but it remains the case that at high levels of utilization, the presence of an additional train on the tracks may lead to additional delays to other trains by reducing the ability of the system to recover from delays. Congestion costs should be distinguished from the costs of delays imposed by the infrastructure manager or by one train operator directly on another. Where these are charged for, they are part of a separate performance regime. Such regimes already exist in Great Britain and a capacity charge based on the cost of congestion has also been introduced (Gibson, 2002).

1.3 Scarcity costs arise where the presence of a train prevents another train from operating, or requires it to take an inferior path. While congestion costs only arise when a train actually operates; scarcity costs are incurred whenever a path is reserved for its use and reflect the characteristics of the foregone service. Quinet (2003) claims that scarcity has not received the attention it deserves. Whilst charging for scarcity is permitted under EU Directive 2001/14, there has been no detailed algorithm for calculating the charge presented.

1.4 It may not generally be appropriate to charge a particular train for both scarcity and congestion, since if a particular train can only be run by displacing another train from the system, it will not be causing additional congestion. However if the congestion caused by the displaced train is taken into account in working out its opportunity cost, then the congestion costs of the train taking its place should still be added to the charge. Given that different trains may cause different amounts of congestion this is actually the most accurate approach.

1.5 The aim of this report is to consider the case for the introduction of scarcity charges as part of rail access charges, the practicalities of doing so and the likely consequences. The next section discusses economic principles, followed by sections on practical experience, identifying bottlenecks, valuing slots, practical implementation issues and likely effects, before we reach our conclusions.

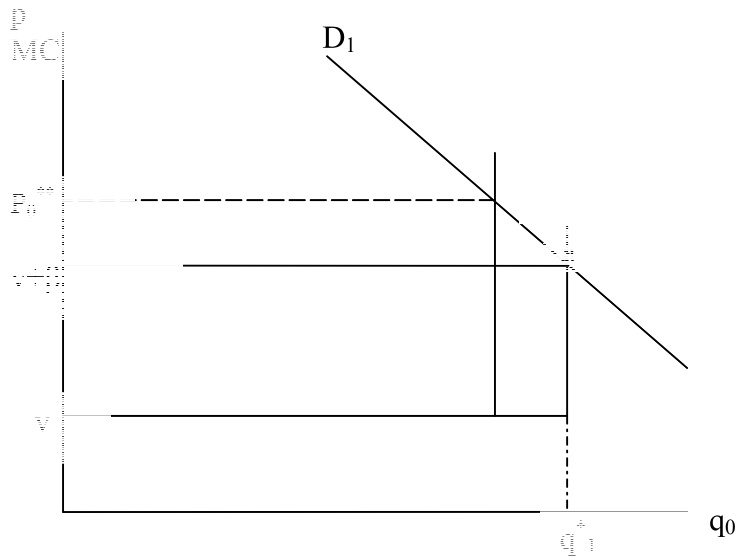
2. Economic principles

2.1 According to economic theory, the most efficient allocation of resources in an economy arises from the pricing of all commodities at marginal social cost. In the case of rail infrastructure this would mean charging train operating companies the wear and tear they cause, any additional congestion costs they impose on other operators, external costs such as environmental or external accident costs and the opportunity cost of the capacity they take up.

2.2 The following analysis is based on Rees (1976). The analysis can be looked at in terms of two periods, period 0 and period 1 or, the short run and the long run for which investments in capacity can be implemented. In year 0, the enterprise, or infrastructure manager, must choose a price and output pair and make an investment plan for year 1, which is derived from the output and price pair planned for year 1. In year 0 output variations can only be made within the limitations of fixed capacity, whereas in year 1 output is fully flexible in line with capacity investment made in year 0.

2.3 The pricing decision is illustrated in figure 1 below. In the diagram, v represents the incremental cost of output when capacity is fixed, or the running costs, β represents the capacity cost of each unit of output, or capital charges. Where there is fixed capacity, the marginal cost pricing rule states that price should be equated to v, the marginal cost, unless there is excess demand at this price, in which case price should be raised to a level to choke off the excess demand to capacity output.

Fig 1 Marginal Cost Pricing and Scarcity



2.4 Suppose in year 0, the demand curve is given by D_1 , and the marginal cost curve is given by MC_0. Note the vertical section of the curve MC_0 shows the point at which capacity is fully utilised. Output cannot be increased beyond this point in year 0 ie the short-run. If price is used to ration the available output, it will have to rise to P_0^{**} generating a profit as this exceeds the running cost and capital charges, $v+\beta$.

2.5 If we now assume that D_1 is the expected demand level in year 1 (and future years), price is now set at the intersection of the new marginal cost curve, MC_1 with the demand curve D_1 at price $v+\beta$ and quantity q^*_1

2.6 In other words, when capacity is optimally adjusted to demand, the scarcity price is exactly equal to the incremental cost of additional capacity. Put another way, this indicates that short run and long run marginal cost are equal when capacity is optimal.

2.7 When they are not equal a choice has to be made between the two alternative approaches to charging. Charging equal to long run marginal cost would give appropriate signals for the long term development of the system, but would not necessarily lead to the optimal use of existing capacity, as the price may be either too high (leading to spare capacity) or too low (leading to excess demand). Given the time lags and indivisibilities in the adjustment of rail capacity this could be a serious problem.

2.8 One option is to charge a variable charge equal to short run marginal social cost as the way of giving correct incentives regarding planning of the next timetable period, but as part of any long run access agreement to charge a fixed charge equal to the avoidable cost of capacity (plus any further financial requirements) which may be renegotiated in the long run if the long run capacity requirements change. Any spare slots could then be sold purely at short run marginal cost, but would imply no long term rights to capacity.

3. Practical experience

3.1 ECMT (2005) presents evidence on the degree to which charges for congestion and scarcity are currently levied in rail track access charges in Europe. Only Great Britain has a congestion charge per train-km explicitly related to estimates of congestion costs.

3.2 A wide variety of both structure (Table 1) and level (fig 2) of charges is found, and it appears there is a range of explanations for this, including differences in the nature and mix of rail traffic, differences in the willingness and ability of governments to provide subsidies, and continued lack of consensus on the measurement of the marginal cost of infrastructure use. The diversity of approaches poses problems particularly for international rail freight, and there is a strong argument for the development of a specific set of international rail freight tariffs.

3.3 When charging for scarcity, it is appropriate to levy a reservation charge, regardless of whether the reserved path is used or not. France has such a charge, whilst Switzerland has a train path cancellation charge which has the same effect. Germany (and the proposals for Slovenia) charges more for ad hoc paths than for regular paths, which is rather the opposite of a reservation fee, but may be justified in terms of costs of train planning.

3.4 With respect to capacity charges, a number of approaches are found and summarised in table 2. Several countries have surcharges on the use of particular stretches of track. For instance, Austria has a surcharge of 0.5 euros per train km for certain bottleneck stretches in the vicinity of Vienna, whilst Denmark has charges of up to 110 euros per train for the use of the key transit routes (this is in addition to bridge charges of up to 900 euros per train for use of the major water crossings; this is a pure cost recovery issue, recovering the railway's share of the costs of these privately financed bridges). Germany also has a surcharge for busy sections. Some countries (Switzerland, Italy) also have charges for the use of busy nodes, and Italy levies a surcharge on trains which is greater the more their speed diverges from the norm for the route in question(whether it is above or below the norm).

3.5 In addition to these explicit capacity charges, some countries (e.g. Germany, France) have separate charges for different categories of route (urban, high speed, secondary etc). Germany distinguishes 11 categories of route and France 12. Generally charges are higher for busier routes, although it is not clear how far this is price differentiation according to demand as opposed to capacity issues.

3.6 Thus it appears that there are a number of countries having something that might be regarded as a scarcity charge, using a variety of approaches. However, we have found no evidence either on how the charge is calculated (it seems to be based on pragmatic judgement rather than any particular formula) or on its effects.

Table 1: Structure of charges

	Pricing Principle	Fixed Charges	Charges per Gross tonne-km	Train-km	Path-km	Other
Austria	MC+		✓	✓		
Bulgaria	MC+			✓		
Czech Republic	MC+		✓	✓		
Denmark	MC+			✓		Charges per train for bottlenecks and bridges
Estonia	FC-	✓	✓	✓		
Finland	MC+		✓			
France	MC+	✓		✓	✓	
Germany	FC-			✓		
Hungary	FC			✓	✓	
Italy	FC- (Traffic management only)	✓		✓	✓	Also charge per node
Latvia	FC			✓		
Netherlands	MC			✓		
Poland	FC			✓	✓	
Portugal	MC			✓		
Romania	FC	✓	✓		✓	
Slovenia	FC			✓		
Sweden	MC+		✓			Oresund bridge surcharge
Switzerland	MC+		✓	✓		Also charge per node
UK	MC+	Franchisees only		✓		Per vehicle km by type of vehicle

Source ECMT (2005)

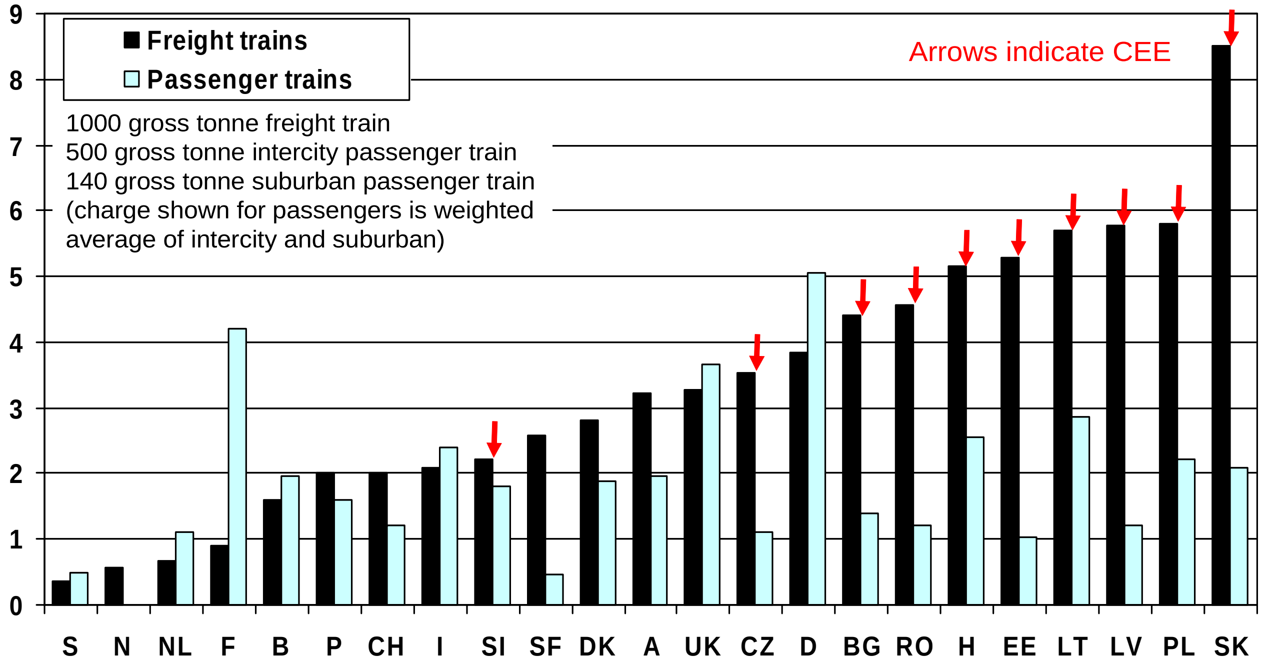
Table 2: Charges for scarce capacity

Reservation fees	Per train path	France
Path cancellation fees	Per train path	Switzerland
Bottleneck charges	Per train km	Austria, Germany
	Per train	Denmark
	Per node	Switzerland, Italy
Differentiation of charges by type of line		Germany, France, Italy*

*also differentiated by speed of train relative to norm for track section

Source ECMT (2005)

Fig 2 Typical levels of charges



Source: ECMT (2005)

4. Identifying Bottleneck sections

4.1 Scarcity charges should be levied only for sections of track where demand exceeds capacity and where this is the key constraint preventing the running of additional trains on a route or routes.

4.2 The carrying capacity of a railway link is the maximum number of physical transport units which can use the link, and can be expressed as a function of the number of tracks in a section, average train speeds, geometry, signalling and safety systems, section lengths, length of trains, etc. (Rothengatter et al, 1996). Moreover, all these factors will vary for individual segments of the route, and the services operating will vary by time of day, requiring calculations to be done in fairly fine detail. However, over and above all these factors, the mix of train speeds and the precise order in which trains are run is crucial. For instance, on a predominantly high-speed line an additional slow freight train may remove the paths of several high-speed passenger trains; on a heavy freight route the reverse may be true. Capacity is also maximised by grouping trains of like speeds, so that a 'flight' of fast passenger trains is followed by a 'flight' of slow freights and vice versa. However, this conflicts with providing a good service of well spaced trains at regular intervals for the public. More complicated still is the interaction of trains on different routes or between different origins and destinations; as with roads, junctions and other bottlenecks (e.g. speed restrictions) are key factors determining capacity.

4.3 The result of all these considerations is that it is impossible to come to a ready definition of the capacity of a rail route corresponding to that for roads. More seriously, the impact of an additional train of a particular type on the paths available to other trains will differ enormously according to the precise mix of traffic on the line. At the same time, the value of a slot to other commercial operators or to government bodies providing social services will also differ enormously in time and space.

4.4 An outcome of work undertaken on congestion for the first periodic review of access charges in the UK (ORR, 2000) was a definition of capacity utilisation. It takes the actual timetable being operated over a period and 'squeezes' the timetable together, accounting for infrastructure / signalling capability until the trains are at the minimum possible operational headway apart, whilst maintaining the original ordering of the trains. Capacity utilisation is defined as the time taken to operate this 'squeezed' timetable as a proportion of time taken to operate the actual timetable. The utilisation index does not explicitly account for interactions between trains at junction bottlenecks. Network Rail suggest that it is generally inefficient to plan for capacity utilisation above 75-80%, because the cost of delays offsets the benefits of running additional trains. Data is published at the national level in the Network Rail Business Plan. As an example, routes with a CUI of over 80% in the morning peak are the main lines to Birmingham from Coventry, Wolverhampton and Longbridge.

4.5 It should be noted that the Capacity Utilisation Index (CUI) was developed for measuring impact of additional trains on secondary delay (congestion). It does not take account of the exact timing of trains (e.g. desirability of regular intervals). Therefore it is likely to overstate practical capacity. Nor does it take account of the many conflicts at stations or junctions which may restrict capacity. Strictly speaking, scarcity charges should be developed separately for individual track sections, junctions and stations, but we doubt whether this level of detail is a realistic proposition. In practice, the impact of running additional trains may not be to displace existing ones, but to lead to inferior patterns of departure times or stops, possibly including the loss of regular-interval timetables.

4.6 Thus it is suggested that the CUI is at best a crude way of identifying problem sections of track. The route utilisation strategies look at these issues in much more detail and identify trade-offs between different uses of paths. These are likely to be a more effective way of identifying the opportunity costs of making additional paths available to particular operators, in terms of what else has to be given up.

4.7 Appendix 2 covers the issue of bottlenecks in more detail and describes a case study for a particular part of the British rail network focussing on the importance of maintaining regular interval services, highlighting specific bottleneck areas and the effect of accommodating freight traffic. It concludes that there is no easy way of identifying bottlenecks given the way in which they interact to constrain timetables; action to remove one bottleneck may simply lead to another one becoming the constraint. It appears that the best that can be done is to adopt a broad brush approach in which all paths which are shown to conflict with the operation of additional services by the main franchisee are charged a scarcity charge which would reflect the full value of the train removed. This would include trains on other routes using platforms or junctions shared with the route in question if these were shown to be crucial to the ability to run additional trains.

5. How might values for slots be established?

5.1 Introduction

5.1.1 This section will consider alternative approaches, including auctioning and the use of models for estimation. It will consider also whether the relevant value is purely the commercial value to the train operating company or the social value taking account of user benefits and externalities. Alternative approaches will be illustrated, making use of ITS's PRAISE model (Preston et al 1999, Johnson and Nash 2005), which simulates demand, revenue and costs for a particular route and timetable, as well as work on valuing benefits of freight traffic.

5.2 Auctioning

5.2.1 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. Auctions would reveal the values of items on which the regulator has an information asymmetry, for example train operator profits. Examples of how such auctioning could be carried out are outlined in the work of Nilsson (2002) and Brewer and Plott (1996).

5.2.2. Nilsson (2002) addresses the issue of auctioning of slots in the context of the EU's drive to charge for scarcity. He observes the following aspects of track access demand:

- Each departure time slot has a large number of substitutes, but each of these may have a different economic value. For freight services, scope for deviations from an ideal slot may be larger.
- There may be cyclical variations creating peak periods where capacity is insufficient
- There maybe complementarities in demand for track access, eg going from A to B may only be valuable if it is possible to proceed to C.
- Services of different operators may be substitutes and operators may value track access differently depending on how close other services operate spatially and temporally.

5.2.3 If trains do not conflict with each other it is a simple enough task to allocate services to operators based on what they wish to run. However, this will rarely be the case - instead, an approach is outlined based on auctioning which combines economic theory with mathematical optimisation to describe how the allocation of slots could be implemented.

5.2.4 The infrastructure manager (IM) will seek to establish a value maximising solution over all operators' values of paths, as at any point in time, there can only be one train, using one and the same 'block' of track. The other constraint is that the schedule must be feasible in terms of the physical limitations on rolling stock and safety. The dual optimisation approach required may not yield a fully optimal solution but will provide a measure of how far from the optimum a specific solution is. The auctioning is carried out as part of an iterative procedure:

- Bidders announce the values for their desired and alternative paths.
- Based on these valuations, the planner optimises total surplus and publishes these results.
- Bidders update their bids, and a new optimisation is carried out until no bidder wishes to change their bid.
- Successful bidders are charged the price bid by the closest unsuccessful bidder which is the opportunity cost of the path.

5.2.5 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 and of the impact on rolling stock requirements). The iterative procedure is intended to deal with these problems, but it would add to the already lengthy process of developing timetables, leading many to doubt its feasibility. 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. The approach itself does not seem particularly suited to dealing with anti-competitive behaviour, where firms can collude to lower the costs they face.

5.2.6 In practice it is far more likely that auctioning could be applied to the allocation of marginal slots, for instance after passenger franchisees have fulfilled their franchise obligations, than to the development of a complete timetable. Even so, it would be necessary to specify timings that did not interfere with allocated slots and to allow for the external costs and benefits of alternative uses of capacity, ideally by appropriate tax and subsidy regimes based on charges and payments per train, passenger or tonne km.

5.3 Use of Models

5.3.1 An alternative to auctioning is the use of models to estimate the value of a slot to a particular train operating company. As a case study we have calculated the value of slots for passenger services using the PRAISE model; and also examined some typical valuations of freight and regional passenger slots using results from other studies (Sansom et al, 2001;Nash, Coulthard and Matthews, 2004). In British circumstances, the Regulator has access to extensive demand and cost data for passenger operations, particularly for franchisees, and it is postulated that this approach is feasible in principle, although further work would be desirable to examine the sensitivity of the results to the precise details of the timetables to be operated and of the alternatives considered.

5.3.2 We use the PRAISE model because it simulates in detail the effect on traffic, revenue and economic impacts of particular changes to the timetable, predicting the actual choice of train of passengers and taking into account the consequent loading of the trains in question. Some exploration of the use of PRAISE to identify opportunity costs was undertaken, but results are not reproduced here due to commercial confidentiality and are specific to particular circumstances considered and assumptions made such as speed and stations served by additional services. It was found that there are large divergences between the social and commercial values of alternative uses of slots, which suggests that neither scarcity pricing nor auctioning slots will lead to an optimal use of capacity unless a way is found of taking these into account. Partly this is because of user and external benefits leading to higher social than commercial values, but also the commercial value of a slot is inflated by the diversion of revenue from other train operating companies; this is not part of the social benefit of a particular use of a slot and thus works in the opposite direction. We find in our case study that at the margin the best use of additional peak slots is by incumbent inter city franchise holders and of off-peak slots by freight operators. Based on our analysis it appears unlikely that regional services would win additional slots on capacity constrained routes.

5.3.3 It should also be noted that whilst passenger trains normally run every day as specified according to the timetable, freight trains run according to demand, and it has been suggested to us that they typically only run on less than 50% of the occasions a slot is reserved for them. This means that the benefits of freight trains need to be scaled down accordingly when comparing the benefits of alternative uses of slots. Of course capacity charges should be regarded as reservation fees. Thus freight operating companies would be charged for slots reserved for them whether they used them or not.

5.4 Other approaches

5.4.1 A third approach, recommended by NERA (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 because 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. Short run marginal cost pricing encourages them to restrict capacity in order to keep price high; whereas a system in which 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, where scarcity charges are levied, to undertake studies to determine the cost of expanding capacity, and to test whether this is justified on cost-benefit grounds. As argued above, we believe that the actual cost of providing capacity is better handled as a fixed charge as part of a long run access agreement.

5.4.2 A further alternative is to simply impose a price and see what happens to demand, and then iterate until demand equals capacity. The risk is, however, that serious distortions may occur whilst the price is adjusting, and that strategic game playing may occur to force the price down by withholding demand, where competition is not strong.

5.4.3 Given the difficulties with all these approaches, it may be thought that the best way of handling the issue is to permit direct negotiation between operators and the infrastructure manager over the price and allocation of slots, including investment in new or upgraded capacity. However, British experience of this approach is that it is complex, time consuming and will not necessarily lead to an optimal outcome given the number of parties involved and the scope for free-riding.

5.5 Conclusion

5.5.1 Whilst we would like to see experiments with the auctioning of spare slots, we believe that the most practical approach to estimation of scarcity charges is the calculation of the opportunity cost of slots using models to estimate the net social benefit provided by alternative use of the slots. This should not be difficult for the incumbent franchisees, where much data is readily available to ORR; it is more difficult for new entrants and particularly freight operators, where specific data on the services desired and flows of traffic served is needed. In what follows, we recommend that charges should be based on the value of a slot to the existing franchisees.

6. How could scarcity charges be implemented in practice?

6.1 There are two key questions that need to be answered before a system of scarcity charges could be implemented. These are:

- how to define a capacity constrained slot ?

- how much to charge for it ?

6.2 On the first point, as commented above, the amount of capacity taken up by a train depends on how its speed compares with the speed of the preceding and following train. Typically, for instance, a freight train on a main line may consume two paths for the prime user of the route, express passenger trains. Generally it would seem sensible to define slots in terms of the predominant user, and to charge other users when they effectively need more than one slot to run a train. In principle, whenever a conflicting movement with another route or service costs a slot on this route (allowing for all other constraints) this should also be charged.

6.3 On the second point, we discussed in the previous section a methodology for determining the opportunity cost of a slot. However, to determine the true opportunity cost requires us to know all possible uses of the slot and to select the best. As has been pointed out, if the Regulator has this knowledge then he may directly determine the optimal use of the slot without recourse to a pricing mechanism (Quinet, 2003), although this would fail to provide a signal to operators to amend the choice of time or route for which they bid.

6.4 What would seem more helpful in the British context would be to regard passenger franchisees as effectively the prime user of the route, having already paid for their reserved use of capacity through the fixed element of the access charge. This could then be seen as a charge which entitled them to the level of access rights specified in their track access agreement. The opportunity cost of a slot would then be estimated as the value of the slot when used by the highest value franchisee. Information for the calculation of this value is generally available as part of the monitoring of franchise arrangements, whilst for possible new entrants it is much more difficult.

6.5 Scarcity charges based on this value might then be levied on freight and open access operators, who currently only pay the variable element of the access charge, and for whom far less information is available to the Regulator. Such a charge would give these operators an incentive to economise in their use of scarce capacity (for instance by changing the time or route of their train) that they do not currently possess.

6.6 For passenger franchisees, the franchise agreement in any event protects them from short term changes in track access charges. However it is also important to signal to franchisees and their funders (especially regional or local government, who are not usually directly involved in decisions about the capacity and capability of the network) the costs of their timetabling decisions when seeking additional access rights. It may therefore be appropriate to charge additional paths sought by franchisees a scarcity charge as well. Obviously if this is the case then the charge should not be so high as to lead to paths being wasted, so it would need to be set somewhat below the value to them of the path. Ideally it would be at the value of the slot to the next best use to the franchisee; there may be evidence on this from the route utilisation strategy exercises. Of course, to avoid discrimination this lower charge would have to apply to all operators.

7. What would be the implications of scarcity charges?

7.1 As will be clear from the above discussion, scarcity charges are not seen as a solution to the general task of achieving an efficient timetable. Franchise agreements and other long term access agreements mean that most paths are effectively allocated for a number of years and route utilization strategy studies guide the longer term allocation of capacity. We see scarcity charges as having their greatest benefit in guiding the allocation of marginal paths which may be available either for franchised passenger operators to increase their services or for freight or open access passenger operators to use.

7.2 From the explorations undertaken using PRAISE, it appears that on a typical inter city route to London, it might be appropriate to levy a scarcity charge of around £20 per train km in the peak and £1 per train km in the off-peak. This would be likely to remove any other services that conflicted with inter city trains in the peak. However, in the off peak, wherever freight traffic at the assumed level of profitability was available it would win any spare paths. Some rather cruder calculations (Nash, Coulthard and Matthews, 2004) suggest that the value per train km of regional services will typically be below the values estimated for both inter city and freight, so that regional services would not be expected to win any spare paths whenever they were in conflict with these sectors for paths.

.

7.3 However two important provisos should be noted. Firstly, such changes would only have the desired result if operators were paid a direct subsidy per passenger or freight tonne km to reflect social benefits. In the absence of such payments, the appropriate charges would be much lower.

7.4 Secondly if the effect of running a particular train is to remove a path for a 300 km inter city service, then the opportunity cost of a slot is the full value of that service – i.e. £6000 on the assumption of £20/train km – and a train only running for part of the route or even crossing it which cost a path for the intercity service should pay the full amount. However the charge should be reduced by the benefit gained by any use of capacity released elsewhere on the route (as where for instance a gap between inter city trains caused by the presence of a suburban train further south permits a regional train to use the flat crossing at Newark).

8. Conclusions

8.1 The existing access charging regime in Britain fails to reflect the opportunity cost of scarce rail capacity. To the extent that prices do not determine the use of capacity, this being determined by franchise agreements, by route utilisation studies and ultimately by the Office of Rail Regulation, this may not be seen as a problem. However, the failure to charge appropriately for scarce capacity may fail to give appropriate incentives to open access and freight operators and to franchised operators seeking extra paths to economise adequately in their use of scarce capacity.

8.2 In the case of franchised passenger operators we consider the most appropriate way to charge for capacity is through the fixed element of the two-part tariff..We believe that it should be possible to negotiate changes in the fixed part of the tariff to reflect additional or reduced costs of capacity as part of long term access agreements. This would give train operating companies and their funders in central or regional government the necessary incentives regarding the use of capacity. A similar approach might be taken with other train operators who acquire access rights under long term access agreements.

8.3 Where marginal slots are available with no long term rights to them allocated as part of a long term contract, then we believe a scarcity charge should be levied. In such cases where a limited number of slots are involved it might be possible to allocate them by auction, although many doubt the practicality of even this limited use of auctions. We would like to see some limited experiments to test this proposition. Otherwise we recommend charging them according to the highest value use by a passenger franchisee. Of course the charge should be a reservation fee, to be paid whether or not the slot is used. This would give freight and open access passenger operators an appropriate incentive for the economical use of slots in terms of time of day and routeing, as well as whether to run the service at all, which currently they do not possess.

8.4 If franchised passenger operators were also charged for additional paths then it would be necessary for the charge to be somewhat below the value to them to avoid paths being wasted (unless it is known that another operator is willing to pay more for the paths). Ideally it would be at the value of the slot to the next best use to the franchisee; there may be evidence on this from the route utilisation strategy exercises. Of course, to avoid discrimination this lower charge would have to apply to all operators.

8.5 In principle charges would be levied for any movement which cost a path for the franchisee, and the opportunity cost of the entire path (i.e. value per train km times train km run) would be charged. However it might be safer to start with a low charge per train km for key routes and to learn from experience with this before moving to more complex structures.

8.6 Scarcity charges can only lead to optimal results if all external social costs and benefits are adequately reflected in charges or subsidies per passenger or freight tonne km. In the absence of this, it appears that much lower charges are generally appropriate.

8.7 On the basis of our case study, we would expect charges on key inter city routes to prevent their use by freight at peak times but that at the margin additional freight traffic would take slots rather than inter city in the off peak. Regional services would be unlikely to win additional slots. However more detailed calculations are needed by route to check the robustness of this conclusion..

9.1 References

AEAT (2002) MOIRA, AEA Technology

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)

Brewer, P.J., Plott, C.R. (1996) 'A binary conflict ascending price mechanism for the decentralized allocation of the right to use railroad tracks', International Journal of Industrial Organisation 14 (6), 857-886.

Department of Transport (1993) 'Gaining access to the Railway network. The Government's proposals'

Department for Transport (2003), 'Transport Economics Note', March

DETR (2001), 'Transport Economics Note', March.

DETR (1998) 'A New Deal for Transport'.

DMRB {??} (2003) Volume 13, Highways Economics Note 1.

ECMT (2005) Charges for the Use of Infrastructure in ECMT Railways. Report and Recommendations. Paris.

Gibson, S, G Cooper and B Ball (2002) 'Capacity charges on the UK rail network.' Journal of Transport Economics and Policy, 36, 2, 341-354

HM Treasury, 'The Green Book', March 2003.

Johnson, D., Shires, D., Nash, C.A., and Tyler, J (2004). 'Forecasting and Appraising the Impact of a Regular Interval Timetable', *European Transport Conference.*

Johnson, D and C. Nash (2005) 'Charging for Scarce Capacity: A Case Study of Britain's East Coast Main Line', Third Rail Conference on Railroad Industry Structure, Competition and Investment, Stockholm

Nash, Chris, Simon Coulthard and Bryan Matthews (2004) 'Rail track charges in Britain – the issue of charging for Capacity'. Transport Policy, 11, 4, 315-328

NERA (1998) , 'An Examination of Rail Infrastructure Charges', NERA, London.

Nilsson, J-E (2002) 'Towards a Welfare Enhancing Process to Manage Railway Infrastructure Access'. Transportation Research A, 36, 5, 419-436.

Office of the Rail Regulator (1994a) 'Framework for the Approval of Railtrack's Track Access Charges for Franchised Passenger Services', A Consultation Document.

Office of the Rail Regulator (1994b) 'Railtrack's track access charges for franchised passenger services. Developing the structure of charges. A Policy Statement.'

Office of the Rail Regulator (1997) 'The periodic review of Railtrack's access charges: a proposed framework and key issues.'

Office of the Rail Regulator, (2000) 'Periodic review of Railtrack's access charges: final conclusions. ORR, London.'

Office of the Rail Regulator (2001) 'Review of freight charging policy: provisional conclusions', ORR, London.

Quinet, E. (2003) ‘Short Term Adjustments in Rail Activity: The Limited Role of Infrastructure Charges’, Transport Policy, Vol.10, No.1.

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, (1) 77-94.

Rees (1976), Public Enterprise Economics, Wiedenfield and Nicholson, London.

Rothengatter W et al (1996) *Bottlenecks in the European Transport Infrastructure.* Technical Report. Study on behalf of the European Centre for Infrastructure Studies (ECIS). April, 1996

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.

Strategic Rail Authority (2003a) ‘Capacity Utilisation Policy: Network Utilisation Strategy’, London: SRA

Strategic Rail Authority (2003b) ‘National Rail Trends. 4th quarter 2002-3’, London: SRA

TAS, ‘Rail Industry Monitor’, (2004)

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.