

December 2011

Using Incentives to Improve Capacity Utilisation

Report for ORR

NERA
Economic Consulting

Project Team

Stuart Holder
Emily Bulman
Helen Smith
Adil Hanif
Peter Spittal

This report sets forth the information required by the terms of NERA's engagement by the Office of Rail Regulation and is prepared in the form expressly required thereby. This report is intended to be read and used as a whole and not in parts. Separation or alteration of any section or page from the main body of this report is expressly forbidden and invalidates this report.

Information furnished by others, upon which all or portions of this report are based, is believed to be reliable but has not been verified. No warranty is given as to the accuracy of such information. Public information and industry and statistical data are from sources we deem to be reliable; however, we make no representation as to the accuracy or completeness of such information and have accepted the information without further verification.

The findings contained in this report may contain predictions based on current data and historical trends. Any such predictions are subject to inherent risks and uncertainties. In particular, actual results could be affected by future events which cannot be predicted or controlled, including, without limitation, changes in business strategies, the development of future products and services, changes in market and industry conditions, the outcome of contingencies, changes in management, changes in law or regulations. NERA accepts no responsibility for actual results or future events.

The opinions expressed in this report are valid only for the purpose stated herein and as of the date of this report. No obligation is assumed to revise this report to reflect changes, events or conditions, which occur subsequent to the date hereof.

All decisions in connection with the implementation or use of advice or recommendations contained in this report are the sole responsibility of the Office of Rail Regulation. This report does not represent investment advice nor does it provide an opinion regarding the fairness of any transaction to any and all parties.

This report is for the exclusive use of the Office of Rail Regulation. There are no third party beneficiaries with respect to this report, and NERA does not accept any liability to any third party. In particular, NERA shall not have any liability to any third party in respect of the contents of this report or any actions taken or decisions made as a consequence of the results, advice or recommendations set forth herein.

NERA Economic Consulting
15 Stratford Place
London W1C 1BE
United Kingdom
Tel: +44 20 7659 8500
Fax: +44 20 7659 8501
www.nera.com

Contents

Executive Summary	i
1. Introduction	1
1.1. Rail Value for Money Study	1
1.2. Capacity Utilisation vs Capacity Allocation	2
2. Existing Regulatory and Contractual Mechanisms	3
3. Capacity Utilisation: The Role of Charges	6
4. Options for a CU Charge	8
4.1. Basic CU Charge	8
4.2. Model-Based Approach	9
5. Stakeholder Views	10
5.1. Consultation Responses	10
5.2. Workshop	10
6. Conclusions	13
Appendix A. Detailed Examples of CU Charging Options	15
Appendix B. Detailed Points Made at Stakeholder Workshop	17
Appendix C. Detailed Questions Raised by ORR	19
Appendix D. Questions Considered at the Stakeholders Workshop	20

Executive Summary

NERA was commissioned by the Office of Rail Regulation (ORR) to consider the scope for using track access charges to provide incentives that will help improve infrastructure capacity utilisation. ORR has previously considered some form of scarcity charge, and the McNulty Rail Value for Money Study identified some features of current service patterns (such as operators not running trains to their full potential length) that may lead to under-utilisation. This report therefore assesses whether there is a role for financial incentives to promote improved utilisation of infrastructure capacity.

Two important considerations are:

- § the significant practical difficulties that would arise if attempting to use charges as a primary means of determining the allocation of train paths on congested parts of the network. Charges would need to reflect the “opportunity cost” of using scarce capacity (ie the value that would have been generated by the best alternative use of the same infrastructure), which is likely to vary significantly between different types of service, different parts of the network and different times of day. Any error in setting charges could well lead to capacity on some of the most valuable parts of the network remaining unused; and
- § the wide range of existing regulatory and contractual mechanisms that already influence the way that capacity is allocated and used, including (among others) ORR’s approval of track access contracts, funders’ specification of franchise and concession contracts, Route Utilisation Strategies and access reviews carried out under condition J of the Network Code.

In view of these, we focused on relatively simple options (though we also considered a more complex bespoke charge that would be tailored to each individual service on the network). These were considered at a small workshop, organised by ORR and attended by invited representatives from Network Rail, passenger and freight train operators, and funders.

A simple option might involve a fixed additional charge per train-km that is applied to each of small number (perhaps 3-6) of network categories, defined by both location and time of day.¹ All services using a given part of the network at a certain time of day/week would pay the same charge. This approach would aim to promote some improvement in capacity utilisation, rather than “solving” the problem and ensuring an optimal allocation of train paths.

One important area of uncertainty is that it is still not clear how much freedom franchised train operators will have to make significant changes to service patterns under the Department for Transport’s revised franchise policy. Even if they have extensive freedom, however, it is not clear that there is significant scope for charges to influence train operators’ decisions in a way that would lead to improved capacity utilisation.

In large part, this reflects a view that infrastructure capacity is already close to being fully utilised in the peak. And where improvements might be possible, these might be best

¹ In one example, categories were London commuting region peak, London commuting region off-peak, other cities peak, inter-urban route peak, inter-urban routes off-peak and other

promoted by specific initiatives (often involving several different parties, and making use of existing regulatory and contractual mechanisms) rather than a general charging mechanism. In addition, there is a risk that a charge aimed at improving capacity utilisation could cause additional distortions, and might lead for example to off-peak services being withdrawn or freight operators switching to alternative routes that contribute to an increase in industry costs. For these reasons, we are doubtful about the scope to improve capacity utilisation by changing train operators' decisions through the introduction of charges or similar incentives.

There could be a useful role for charges, however, in providing stronger incentives for Network Rail to be more proactive in taking measures to improve capacity utilisation (for example, identifying possible timetable changes that would free up additional train paths). At present, Network Rail (and its managers) faces strong incentives to improve operational performance, and these might discourage it from taking action that could increase the number of services using congested parts of the network. Incentives targeted on Network Rail could be introduced in a way that would not directly affect train operators' decisions. But it would be important to carry out further analysis, both to assess the likely effectiveness of such measures, and also to consider any risks of distortions or unintended consequences.

1. Introduction

This report, prepared by NERA for the Office of Rail Regulation (ORR), considers the scope for using incentives determined at a periodic review, in particular the track access charging framework, to help improve infrastructure capacity utilisation. It provides a general assessment of the potential role for financial incentives in promoting improved capacity utilisation, and also provides some examples of specific charging options. Our analysis and conclusions also reflect a discussion of these issues at a small stakeholder workshop on 23 September. We would like to thank those who participated in this workshop.

The rest of this section provides some important background, before Sections 2 and 3 describe some existing regulatory and contractual mechanisms and discuss the potential role for charging mechanisms. Section 4 then sets out some illustrative options and Section 5 summarises the main points made by stakeholders in responding to ORR's first consultation on PR13 and at the workshop. Section 6 contains our conclusions.

In this report, we use "CU charge" as a generic name for a charge that is designed to promote a better use of capacity. It may differ from the type of scarcity charge previously considered by ORR. It can also be distinguished from the existing capacity charge, which is intended to compensate Network Rail for any expected increase in Schedule 8 payments (see section 2), and also Network Rail's volume incentive which applies in the same way to both congested and uncongested parts of the network.

1.1. Rail Value for Money Study

The Rail Value for Money Study examined capacity management and utilisation,² concluding that "if the industry can change its mindset regarding the utilisation of existing system capacity in the coming years and, in particular, focus on improving average loadings of existing, more lightly loaded services, there is an opportunity (with projected passenger growth and a slowly reducing cost base) to significantly improve unit costs in the industry". While the international comparisons cited in the study can be interpreted in a number of different ways, not all of which suggest an inefficient use of capacity, the report also identified a number of specific features of current service patterns which may lead to under-utilisation. These include:

- § heavily-peaked demand, with the result that "off-peak trains are invariably not fully-loaded";
- § not running trains to their full potential length, often reflecting commercial incentives or franchise requirements to run more frequent, shorter trains; and
- § a large number of "through-journey opportunities", leading to short or poorly-loaded trains taking up capacity on congested parts of the network.

² See Section 17 of the Final Report, "Realising the potential of GB rail: final independent report of the Rail Value for Money study", May 2011, <http://www.rail-reg.gov.uk/upload/pdf/rail-vfm-detailed-report-may11.pdf>.

1.2. Capacity Utilisation vs Capacity Allocation

Making the most effective use of the existing infrastructure capacity can be considered as two related questions. The first is whether capacity is utilised effectively, for example, whether the maximum number of train paths has been made available using the existing infrastructure. If no additional train paths can be found to accommodate proposed new services, then an important second question is whether the existing capacity is being allocated to the type of services that make the best use of it (either because they generate the most revenue, or because they generate significant user benefits or external impacts, such as reducing road congestion).

Each of these questions, in turn, can be considered from different perspectives. Capacity utilisation may be improved if train operators adjust their service patterns in order to use capacity more efficiently (for example, running fewer, longer trains), thereby freeing up capacity that can now be used by other services. And capacity utilisation may also be improved if Network Rail is more proactive in identifying potential train paths, either because it takes a less risk averse attitude to potential performance degradations, or because it increases its efforts to identify timetable or other changes that could allow extra services to be accommodated.

With regard to capacity allocation, one important question is whether any capacity that is (or becomes) available is allocated to the train operator that will make the best use of it. Individual train operators may have very different ideas about the types of service that could be provided. A more far reaching (and controversial) question is whether there might be scope to reallocate existing access rights if the current mix of services is suboptimal.

The next section describes some of the other regulatory and contractual mechanisms that may affect capacity utilisation and allocation. Charges that aim to improve capacity utilisation might be introduced either instead of, or alongside, some of these mechanisms.

2. Existing Regulatory and Contractual Mechanisms

There are a number of existing regulatory and contractual mechanisms that already influence the way that capacity is allocated and used. It will be important, therefore, to consider whether a potential CU charge might operate either alongside or instead of these mechanisms. This in turn raises questions of whether charges are likely to lead to better outcomes than some of the existing mechanisms, and whether certain existing mechanisms might need to be adapted.

The amount of capacity available, and whether measures are taken to relieve capacity bottlenecks, are significantly affected by the periodic review process, including the High Level Output Specification (HLOS) and Statement of Funds Available (SoFA) which define the capacity the government is willing to buy.

The way that this capacity is then allocated and used will be directly affected by mechanisms that include:

- § ORR's approval of track access contracts. This provides an opportunity for ORR to address situations where the definition of access rights might restrict Network Rail's ability to make the best possible use existing capacity. In cases where not all proposed services can be accommodated on the existing network, it then allows ORR to decide between alternative applications;³
- § ORR's policy on on-rail competition, including the current "not primarily abstractive" test, which has an important influence on the type of services that open access operators can provide in competition with TOCs;
- § the specification of franchise and concession contracts by Department for Transport (DfT), Transport Scotland (TS), TfL and MerseyTravel,⁴ which until recently has tended to involve a rather detailed definition of the specific services to be provided by franchisees. This reduces the scope for train operating companies (TOCs) to propose alternative ways of delivering the service (for example, longer but less-frequent trains) that might free up additional capacity;
- § Route Utilisation Strategies (RUSs) which, in addition to identifying possible options for providing additional capacity, may provide (non-binding) suggestions on how existing capacity can be better used (for example by optimising timetables), and how (if at all) expected future demand for additional train paths can be accommodated. RUSs have now been produced that cover the entire rail network;
- § Network Rail's Code of Practice on the Management of Strategic Capacity on the Network, which sets out how Network Rail will identify available capacity on key routes

³ ORR publishes the criteria and procedures it uses to consider track access applications. In the event of competing claims on the available capacity, it may apply cost-benefit analysis to inform its decision on the most appropriate allocation.

⁴ Transport for London (TfL) has responsibility for the London Overground concession.

to create Strategic Paths, include these paths in the timetable development process, and maintain them as far as practicable during the development of future timetables;⁵

- § access right reviews carried out by Network Rail and other “use it or lose it” mechanisms, as set out in Part J of the Network Code, which may provide an opportunity to challenge access rights that are inefficiently used. Historic access rights, for example for some coal traffic, may have been established at a time when capacity was less of an issue;
- § Network Rail’s volume incentive, which provides a financial reward to Network Rail if growth in passenger and freight traffic and passenger revenues is higher than the baseline set by ORR during the last periodic review. It therefore provides an incentive for Network Rail to make more capacity available to prospective users;
- § track access charges overall, and specifically the existing capacity charge, which is intended to compensate Network Rail for the expected increase in Schedule 8 payments because of the impact on reactive delays from accepting new traffic.⁶ It therefore aims to remove the disincentive to accept additional traffic that Network Rail would otherwise face; and
- § performance targets for Network Rail set by ORR and monitored on an ongoing basis.⁷ Notwithstanding other incentives, the risk of increased delays could discourage Network Rail from making capacity available on parts of the network that are already congested.

Table 2.1 identifies the main mechanisms that affect different aspects of capacity utilisation and allocation at present. A tick (ü) indicates a mechanism that may be likely to promote improvements in specific aspects of capacity utilisation and allocation, whereas a cross (û) indicates a mechanism that might frustrate potential improvements, for example by blunting incentives. ORR’s policy on on-rail competition will affect decisions about the allocation of available train paths. However, because of the “not primarily abstractive” rule it is not clear that this will necessarily lead to optimal use of capacity, as it may increase the number of services that run through popular stations without stopping. The outcome reflects, among other things, the result of trade-offs between ORR’s statutory duties. Similarly, while DfT and other funders may carry out some analysis of alternative options during the process of drafting franchise specifications and selecting preferred bidders, the scale and scope of such analysis (for example the extent to which it takes account of potential new freight or non-franchised passenger services) may fall short of that necessary to ensure the optimal use of the available capacity.

⁵ Network Rail (2011), “Management of Strategic Capacity on the Network: Network Rail’s Code of Practice, Version 1.1”.

⁶ Although when the charge was introduced, it was intended to be implemented in a disaggregate way, with higher charges for peak times and route sections, in practice charges are differentiated only by service group, with weekend discounts. As a result, in very congested parts of the network, the capacity charge may not reflect Network Rail’s extra Schedule 8 payments in full, and this may affect Network Rail’s incentives to provide additional capacity. Conversely, on other sections of route the capacity charge may significantly overestimate the extra schedule 8 payments that Network Rail is likely to make. Given other pressures on Network Rail to improve performance, moreover, the capacity charge may not adequately offset the disincentive to improve capacity utilisation because of the risk of extra delay.

⁷ On routes where performance is significantly behind target, ORR meets with Network Rail to discuss what it is doing to recover the position. In some cases, ORR will then require Network Rail to produce more robust improvement plans. Ultimately, ORR could take formal enforcement action against Network Rail for a breach of its licence.

Table 2.1
Charges, policies and administrative mechanisms

	Operators reconsider their use of existing train paths	Network Rail identifies additional train paths	Optimal allocation of available train paths
ORR approval of track access contracts	-	-	ü
DfT/TS specification of franchise contracts	ü	-	?
ORR's policy on on-rail competition	-	-	?
Route Utilisation Strategies	ü	ü	ü
Management of Strategic Capacity	-	ü	-
Network Rail access right reviews	ü	ü	-
Network Rail's volume incentive	-	ü	ü
Capacity charge	ü	ü	-
Performance targets and monitoring	-	ü	-

Source: NERA

Some changes to this situation are likely in the short to medium term. DfT has announced its intention to move to longer duration franchises with looser specifications than at present. However, it is not yet clear how much flexibility TOCs will have in practice.⁸ Future franchising policy in Scotland is also likely to change, and Transport Scotland is consulting on several changes in its approach to franchising including a shorter franchise, and a dual-focus franchise covering highly specified social rail, and less specified economic rail.⁹

Following the Rail Value for Money study, moreover, some form of demand risk-sharing between Network Rail and train operators may be implemented – potentially as part of an alliance or joint venture. This could significantly strengthen Network Rail's incentives both to identify additional train paths and to ensure that the available capacity is allocated to the train operators that can make best use of it.

Closer co-operation between Network Rail and train operators could, alongside the High Level Output Specification (HLOS) process, also affect the amount of capacity that is provided in future. And any other changes to the existing structure of charges may also affect the incentives on train operators or Network Rail to improve the use of the network.

⁸ The draft ITT for the West Coast Mainline franchise gives a possible indication of the principles for franchising of intercity routes (though DfT may adopt a different approach for London commuter or regional franchises). This suggests that TOCs will have more flexibility to change the number of station stops on different days of the week, but the overall scope to reduce the number of services from the number offered in the current timetable appears to be limited.

⁹ Transport Scotland (2011), "Rail 2014 – Public Consultation", <http://www.transportscotland.gov.uk/strategy-and-research/publications-and-consultations/j203179-00.htm>.

3. Capacity Utilisation: The Role of Charges

Scarce resources in other industries (eg spectrum rights) have been allocated by market mechanisms (principally auctions), and policy makers have considered the potential use of higher prices or auctions to take-off and landing slots at congested airports. However, we do not believe that market mechanisms can be used as a *primary* means of determining the allocation of train paths on congested parts of the network. Key reasons for this include the strong likelihood that different potential users (eg freight and passenger services, or services with different stopping patterns) will require different types of train paths, the very substantial complications involved in trying to set a market clearing set of prices (or bidding in auctions), and the very high risk that errors might lead to capacity on some of the most valuable parts of the network remaining unused. And if there are differences between the economic value of services and the private commercial value to train operators then, absent other measures, the most economically valuable services may be priced off the network at market clearing prices if there are services that generate higher commercial returns.

A more realistic objective, therefore, is to adjust the structure of charges so that it will promote some improvement in capacity utilisation. This may be through providing incentives for train operators (and funders) to economise on their use of capacity, or for Network Rail to identify ways of providing additional train paths. Or it may provide an initial “filter” to eliminate some (but not all) competing applications for available train paths. This could lead to reduced administration costs, and also reduce the risk that administrative decisions will lead to train paths being allocated to services that use capacity inefficiently.

An important concept when considering the appropriate level of charges is the “opportunity cost” of using scarce capacity. This is the value that would have been generated by the best alternative use of the same infrastructure (such as another train service that was unable to run). It therefore measures the value foregone as a result of allowing a particular service access to a congested facility.¹⁰

An ideal CU charge would reflect the opportunity cost of each service at each different location on the network. An added advantage of this would be greater transparency, which might also help facilitate the Rail Value for Money Study’s recommendation that passengers and funders should pay their full marginal cost (including opportunity cost). Even a simpler CU charge should be based on the opportunity cost of using different parts of the network at different times of the day or week.¹¹ An obvious difficulty, however, is that opportunity costs cannot be observed.

¹⁰ Even if some potential train paths remain unused, the opportunity cost will still reflect factors such as reduced operational resilience (which includes the impact on all parties, and is therefore higher than the impact on Network Rail’s Schedule 8 payments that the capacity charge is intended to cover) and the impact of restrictions on timetabling/access rights that operators have accepted in order to create spare capacity.

¹¹ In theory, the ideal charge would be based on a bespoke calculation for each set of services, reflecting the amount of capacity they take up and the resulting opportunity cost. This would take into account local circumstances, such as the other services running on the same part of the network and their value. However, this is infeasible for various reasons including the amount of data that would be required for all services and computational complexity. In addition, the calculation would be complicated by interactions between different services, as the opportunity cost of one service would depend on which other services might want to operate on the same route.

One practical solution is to base a CU charge on some measure of capacity utilisation. While a Capacity Utilisation Index has been developed in the UK, there are many practical problems associated with the definition and measurement of rail capacity. Even if the index indicates that capacity is available, it may not be possible to accommodate the specific additional services (defined by calling patterns, journey times, etc) that operators would like to run.

A more general problem is that there is unlikely to be a perfect correlation between capacity utilisation and opportunity cost. Nevertheless, in view of all the problems noted above, we believe the relevant question is whether a CU charge can *improve* capacity utilisation, rather than seeking a charge that aims in isolation to optimise capacity utilisation and allocation (which would be very complex to implement and carry a high risk of unintentionally pricing traffic off valuable parts of the network). Such a charge will not “solve” the problem of capacity utilisation and allocation. Instead, the question to be considered is whether it might offer some improvement, either instead of or alongside existing administrative mechanisms without adding unnecessary complexity. In this context, a measure of capacity utilisation may provide an appropriate starting point for considering how and where a charge might be applied.

Other possible changes to the structure of charges, such as the use of long run incremental costs (LRIC),¹² have been suggested which might also lead to an increase in variable charges and therefore provide stronger incentives for train operators and Network Rail to improve capacity utilisation. However, the rationale for introducing such charges is different and, importantly, they would also apply to uncongested parts of the network and times of day. In this note, we focus on changes to the charging structure that are designed specifically to promote improvements in capacity utilisation. LRIC-based charges have also been suggested as a means of providing incentives for Network Rail to undertake small projects to increase capacity – this is beyond the scope of this note, and in any case we have strong reservations about whether variable charges can be used to provide effective investment incentives.

¹² LRIC is used extensively to determine charges in telecommunications and some other industries. This is often in the context of vertically integrated companies (though some may be subject to accounting separation or ringfencing requirements) supplying network services to their direct competitors. In such cases, LRIC provides an appropriate basis for seeking contributions to the overall cost of the network (rather than setting interconnection or similar charges at short run marginal cost). But such considerations do not apply to Network Rail’s track access charges, as it does not operate trains and it has also access to an income stream (fixed charges and Network Grant) that can cover its non-variable costs.

4. Options for a CU Charge

In broad terms, we believe there are two main ways that a CU charge (ie a charge specifically aimed at promoting a better use of capacity) could be implemented. The first approach is a fairly basic charge, dividing the network into a small number of categories based on a broad measure of capacity utilisation. The second approach is a more detailed CU charge – because of the way that opportunity costs vary between different services and locations, we believe this would need to be based on a bespoke estimate of the opportunity cost associated with each service. Consistent with the objective of reflecting opportunity costs, these charges would be levied on train paths that are booked, rather than trains that actually run.

Both of these options were presented to and discussed at a small stakeholder workshop organised by ORR in September 2011.

4.1. Basic CU Charge

A basic CU charge would apply to a small number of categories, defined based on capacity utilisation. The categories would be defined by geography and time of day, rather than type of service, and charges would be paid by all passenger and freight trains using a particular route. Appendix A provides two examples of how this approach might work in practice.

The simplest approach (Option 1A) is a CU charge for three categories: low, medium and high utilisation. Charges could be set to zero for sections of the network with low capacity utilisation, and a single rate per train-km for high utilisation sections/times of day. The three sub-options presented at the stakeholder workshop had rates of between £5, £15 or £30 per train-km for the high utilisation category, and £1, £5 or £10 per train-km for the medium utilisation category. For comparison, we note that the average level of the variable usage charge and capacity charge (combined) is somewhat less than £2 per train-km. ITS' 2006 modelling suggested that, for 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.

Option 1B is a similar approach but includes a greater number of charging categories. This would allow the charge to differentiate more between different levels of capacity utilisation. The charging categories adopted for this example are London commuting peak, London commuting off peak, other cities peak, intercity peak, intercity off peak and other, though clearly other categorisations are possible. By adopting narrower categories and more differentiated charges, this approach could allow a more narrowly defined high utilisation category and perhaps higher charges. For the stakeholder workshop, however, we focused on disaggregating the medium and low utilisation routes/times into a greater number of categories.

The basic CU charge approach is simple and aims to reflect the general tendency for high opportunity costs where the network is busy and train paths may be either unavailable or unsuitable for the kind of service that an operator would like to run. The charges are not intended to choke off all excess demand or ensure that capacity is allocated to the most appropriate users. They will not attempt to reflect differences in capacity utilisation along a route, or to reflect problems associated with a specific bottleneck. Nevertheless, by levying additional charges for the most congested parts of the network, this might encourage train

operators to reconsider the way that they use capacity (eg running longer passenger trains, or changing the timing of freight trains), provide useful price signals for funders, and also provide stronger incentives for Network Rail to ensure that as many train paths as possible are made available.

The design of a CU charge will affect its ability to change incentives, for example the number of charging categories. A CU charge with few charging categories (such as Option 1A) has the advantages of being simple to design and implement, but it is possible that such an aggregate system of charges would not address capacity utilisation incentives in a sufficiently focused way, and could distort incentives. Having a greater number of categories (as in Option 1B) provides a more disaggregate approach, which may allow the charge to be more focused on particular capacity utilisation problems in different parts of the network.

Even with a greater number of charging categories, however, it may still be difficult to set the correct relationship between different charges without creating perverse incentives. For example, different charges might need to be set for peak and off peak charges on London commuter routes. Setting the peak charge too high relative to the off peak charge might drive services out of the peak and simply create further capacity utilisation problems off peak.

4.2. Model-Based Approach

Another approach would be to use a model to set a bespoke CU charge, using inputs for existing services, the available capacity at different points along the route and the nature of the proposed new service, plus pre-defined model parameters. The same model would be used for each train, but it would be run using different detailed data inputs (including the nature of the new service, and details of existing services using the same route at the same time of day) and would therefore generate a specific CU charge for each service. A study for ORR by the Institute for Transport Studies (ITS) provides one example of how a modelling-based approach might work.¹³

A model-based approach would result in the setting of CU charges more closely aligned to opportunity costs, allowing incentives to be focused on areas with capacity utilisation problems. As this would also take into account existing services, it could well lead to very different charges applying to services (eg non-stop passenger services, stopping passenger services or freight services) wishing to use the same route at the same time of day as each other. For this reason, we would expect a higher charge to be set for a proposed service that is difficult to timetable, for example because of a disparity between the average line speeds of existing and proposed new services. Some charges for specific bottlenecks and “difficult” trains could well be higher than the £30 per train-km used for illustrative purposes above.

¹³ Institute for Transport Studies, *Scoping Study for Scarcity Charges: Final Report for the Office of Rail Regulation*, December 2005 (revised February and March 2006).

5. Stakeholder Views

5.1. Consultation Responses

ORR's first consultation on PR13 asked stakeholders to consider (among many other things) "how the effectiveness of Network Rail's incentives to make best use of capacity could be improved". A more detailed list of questions was included in Annex D to ORR's May 2011 consultation (these are reproduced in Appendix C).¹⁴

Some consultation respondents supported the idea of a scarcity charge, though concerns were expressed about increasing the complexity of the existing structure of charges. Some respondents were also sceptical about whether train operators would be able to respond to any new incentives in practice (and therefore it might become little more than a tax). Other respondents did not support a scarcity charge, referring also to the availability of non-financial mechanisms to promote improved capacity utilisation and the fact that no similar charge applies to road hauliers using the national road network.

There was also some confusion about whether, and how, a scarcity charge would be financially neutral. Some respondents appeared to expect a reduction in off-peak charges. Others argued for financial neutrality, but nevertheless expressed concerns about the impact of such a charge on particular traffic flows.

5.2. Workshop

As part of the current project, and in order to obtain input from industry participants on the potential impact of different options for incentivising improved capacity utilisation, ORR organised a small workshop attended by invited representatives from Network Rail, passenger and freight train operators, and funders. NERA presented both the general principles set out in previous sections and also the illustrative options described in Section 4 and Appendix A. A constructive discussion followed. The detailed questions that participants were asked to consider are reproduced in Appendix D, and a list of some of the specific comments made at the workshop is at Appendix B.

Discussion at the workshop focussed almost entirely on the possible impact of a basic CU charge. No participants expressed any enthusiasm for a modelling-based approach (which would also cause particular problems for freight operators, who would have to quote charges to customers before they know what track access charges they would have to pay). And some stakeholders were cautious about any change to the charging framework, which they view as already too complex.

An important area of uncertainty is that it is still not clear how much freedom TOCs will have to make significant changes to service patterns under the DfT's revised franchising policy. Even if new franchise specifications do permit significant changes, many train operators were sceptical about whether a CU charge would improve capacity utilisation. In large part, this reflected doubts about whether any improvement was possible, irrespective of the financial incentives available. Especially in the peak, there was a widely-held view that capacity was

¹⁴ ORR (2011), "Periodic review 2013: first consultation – annexes", paragraph D60

at least close to being fully utilised. And where this was not the case, it often reflected the impact of other constraints, such as:

- § a lack of rolling stock in cases where some trains fell short of the maximum possible length;
- § a lack of funding to cover the costs of additional or enhanced services; or
- § the need for well-defined track access rights, so TOCs could be confident that they will have the train paths they need, but which can make it more difficult to change service patterns.

Furthermore, several participants argued that, where opportunities to improve capacity utilisation might exist, co-operation between many different industry parties was required to implement such improvements. Participants did not believe the CU charge would provide incentives for such co-operation, and this applies especially in cases where a train operator might take actions to free up capacity which could then be exploited by another train operator (and perhaps a direct competitor).

In addition to general doubts about the impact of a CU charge on train operators' decisions, there was concern about possible distortions that a CU charge might cause. For example:

- § off-peak services could be priced off the network, if train operators look to make savings in response to higher charges, but peak services are still the most profitable to them (even after the CU charge). As well as leading to unnecessary service cuts, this could also frustrate other objectives, such as encouraging passengers to switch from peak to off-peak services;
- § freight traffic could transfer to road, if customers want trains to run at peak times but the CU charge means that this is too expensive;
- § freight traffic could also divert to less congested routes, which may not be efficient from an industry-wide perspective (for example, because it will lead to an increase in infrastructure maintenance costs);
- § the impact of higher costs on some funders could lead to service cuts (probably off-peak) because of affordability constraints.

A CU charge might need to be relatively large in order to affect train operators' decisions in the way intended, but this would also increase the risk of the types of distortion listed above.

Some stakeholders felt that a CU charge would not achieve a meaningful impact unless there were significant changes in wider policies (about the most appropriate use of congested infrastructure). There was unease, moreover, about the risk of a CU charge leading to unintentional policy changes, and some support for looking to improve the RUS process or dealing with any problems arising on a case-by-case basis (especially if most opportunities for improvement are likely to fall outside of the peak).

In contrast, views about the possible impact on Network Rail's behaviour were more positive. This reflected two main factors. The first was a recognition that there may be cases where Network Rail can find ways to change timetables so as to free up additional capacity. The recent case of the finding of an additional path on the East Coast Main Line provides one

example of this, but also highlighted a risk that Network Rail may not be sufficiently proactive in identifying such opportunities.

The second factor was a concern that Network Rail faces a number of strong incentives, on both the company and its managers, to improve operational performance (delays and cancellations). Since accommodating additional traffic on already congested parts of the network is likely to increase expected reactive delays, and notwithstanding the existing capacity charge (which aims to compensate Network Rail for the expected Schedule 8 impact), the risk of a deterioration in performance could well discourage Network Rail from taking proactive steps to identify possible extra train paths.

Several stakeholders could see potential benefits, therefore, if a CU charge helped in some way to redress the apparent imbalance in Network Rail's incentives. But the need for caution was also stressed, because of the risk that strong incentives could lead to further distortions or unintended consequences.

6. Conclusions

For the reasons set out in Section 3, we do not believe that financial incentives should be considered as a primary means of determining the allocation of train paths. So the question we have considered here is whether the structure of charges might be adjusted so as to promote an improvement in capacity utilisation. We have also given examples of simple approaches that might be relatively straightforward to implement, consistent with this more realistic objective.

Despite the comments in the Rail Value for Money Study, however, it is not clear that there are widespread problems of poor capacity utilisation caused by train operators' decisions (and which therefore might be potentially addressed by some form of CU charge). Even if some problems can be identified:

- § they may well be specific to particular locations or circumstances. These problems might not be solved by a general CU charge designed to improve capacity utilisation across the network as a whole;
- § in contrast, there are a number of alternative regulatory, contractual and planning mechanisms, many of which are suitable for addressing specific problems on a case-by-case basis.

Even if examples of poor capacity utilisation were more widespread, it would be important to consider whether improvements to some of these existing mechanisms might provide a more effective solution than some form of CU charge or a similar financial incentive mechanism. And it is quite possible that, where capacity utilisation is sub-optimal, this could reflect the impact of other constraints (such as services that are mandated by funders, shortage of rolling stock, lack of funding for new services, etc) or deliberate policy decisions (for example to provide services to regions, or to promote competition), and therefore unlikely to be resolved by new incentive mechanisms.

Equally, it is not clear what the impact of the DfT and Transport Scotland's new franchising policies will be. If, in practice, franchise specifications are still relatively prescriptive, or the impact of other constraints (such as train paths that other operators use, or the physical and operational characteristics of the route) mean that large scale timetable changes are not possible or are unrealistic commercially, then a CU charge might have little or no impact on train operators' behaviour.

Against this low probability of achieving improvements by changing train operators' behaviour, there remains a non-negligible risk that a CU charge would introduce additional distortions, perhaps to the extent that some services might be unintentionally priced off the network, and would add further to the complexity of the charging and incentive framework. It would be important, therefore, to carry out a more detailed assessment of the likely impact of any proposed charge, in order to provide a more accurate assessment of the likely benefits and costs, and the trade-off between accuracy and simplicity. On balance, therefore, we are doubtful about the scope to improve capacity utilisation by changing train operators' decisions through the introduction of charges or similar incentives.

The arguments for providing incentives for Network Rail to be more proactive in taking measures to improve capacity utilisation appear stronger. In order to avoid the risk of distorting train operators' decisions or pricing traffic off the network, such incentives could be implemented in a similar way to the existing volume incentive. Indeed, a change to the current volume incentive, so that incentives to increase traffic are skewed towards the most congested parts of the network, might be one way of implementing such an approach. Before doing this, however, it might be useful to gather more information about the effectiveness of the current incentive, and whether there are ways it could be made more effective, or whether there are risks of distortions. There would also be a number of detailed questions to be addressed about how a more targeted CU incentive would be implemented.

A final important caveat is that this paper has addressed the question of whether financial incentives could be used to improve capacity utilisation. It has focused, moreover, on ways in which the rail network might be used sub-optimally, rather than considering other aspects of capacity utilisation (such as loading factors on trains).

It is possible that a CU charge, or a similar mechanism introduced with the aim of improving capacity utilisation, might bring other potential benefits. For example, more general arguments have been advanced for changes that would lead to a higher level of variable charges, for reasons such as making Network Rail more responsive to train operators' needs or providing more useful price signals to funders. A CU charge might meet some of these objectives (irrespective of whether or not it improves capacity utilisation). But there could well be more effective ways of providing such benefits, including mechanisms that are designed specifically for these purposes.

Appendix A. Detailed Examples of CU Charging Options

The following examples of a basic CU charge (see Section 4.1) were presented to the stakeholder workshop.

Box A.1 Option 1A: Basic Price List

Option 1 involves different CU charges per train-km for a small number of utilisation categories. Sections of the network would be divided into the utilisation categories similar to those below, based on a capacity utilisation metric.

- § **High utilisation** – perhaps including weekday morning and evening peaks for London commuter region (eg area bounded by Reading, High Wycombe, Milton Keynes, Bedford, Hitchin, Bishop’s Stortford, Shenfield, Southend, Tonbridge, Gatwick and Woking)
- § **Medium utilisation** – perhaps including (a) off-peak for London commuter region (as defined above), (b) weekday morning and peaks for Birmingham, Liverpool, Manchester, Leeds, Newcastle and Glasgow areas, (c) weekday daytime services on mainline routes between London and Plymouth, Bristol, Cardiff, Birmingham, Liverpool, Manchester, Glasgow, Derby, Nottingham, Sheffield, Leeds, Edinburgh, Cambridge, Norwich, Dover, Brighton and Southampton (but excluding the London commuter region, which belongs to the high utilisation category)
- § **Low utilisation** – all other routes and times of day/week.

Table A.1
Option 1A: Illustrative Charges

£ per train-km	Option A1	Option A2	Option A3
High utilisation	5	15	30
Medium utilisation	1	5	10
Low utilisation	0	0	0

Box A.2
Option 1B: More Detailed Price List

Option 2 would have a larger number of categories than Option 1, to allow a more focused impact on incentives. Option 2 would have different charges per train-km for the following categories. Sections of track would be assigned to a charging category based on the predominant use of the section.

- § **London commuting peak** - weekday morning and evening peaks for London commuter region (eg area bounded by Reading, High Wycombe, Milton Keynes, Bedford, Hitchin, Bishop’s Stortford, Shenfield, Southend, Tonbridge, Gatwick and Woking)
- § **London commuting off peak** – off-peak services in London commuter region (as defined above)
- § **Other cities peak** - weekday morning and peaks for Birmingham, Liverpool, Manchester, Leeds, Newcastle and Glasgow areas
- § **Intercity peak** - weekday daytime services on mainline routes between London and Plymouth, Bristol, Cardiff, Birmingham, Liverpool, Manchester, Glasgow, Derby, Nottingham, Sheffield, Leeds, Edinburgh, Cambridge, Norwich, Dover, Brighton and Southampton (but excluding the London commuter region, which belongs to a separate category)
- § **Intercity off peak** - evening and weekend services on mainline routes between London and Plymouth, Bristol, Cardiff, Birmingham, Liverpool, Manchester, Glasgow, Derby, Nottingham, Sheffield, Leeds, Edinburgh, Cambridge, Norwich, Dover, Brighton and Southampton (but excluding the London commuter region, which belongs to a separate category)
- § **Other** - all other routes and times of day/week.

Table A.2
Option 1B: Illustrative Charges

£ per train-km	Option B1	Option B2	Option B3
London commuting peak	5	15	30
London commuting off peak	1	5	10
Other cities peak	2	10	20
Intercity peak	2	10	15
Intercity off peak	0	1	5
Other	0	0	0

Appendix B. Detailed Points Made at Stakeholder Workshop

B.1. Effects of CU Charge on Behaviour of TOCs and Network Rail

- § Altering variable charges would not affect behaviour without there also being policy changes (eg franchising policy).
- § The scope for lengthening trains is already limited and within five years further lengthening is likely to be impossible.
- § A CU charge might lead to better use of the network but lower total demand (as a result of a reduction in frequency), and therefore more lightly loaded trains.
- § A CU charge should provide an incentive to move traffic onto the shoulders of the peak.
- § Revenues are a larger driver of TOC behaviour than charges.
- § If the CU charge was significant enough to change behaviours, it would represent a subtle shift away from an economic rationale for service provision towards a financial rationale.
- § While capacity utilisation could be improved through cooperation between TOCs, a CU charge would not incentivise TOCs to work together, which would be the most likely way to find extra capacity.
- § TOCs and Network Rail do not have a long-term interest in profit so may be less responsive to financial incentives.
- § TOCs will still need good access contracts with Network Rail even if franchise specifications are relaxed, so may not have much flexibility in practice.
- § A CU charge might cause Network Rail to favour enhancements over timetable changes.
- § There is an overemphasis on performance in RUSs as a result of pressures on Network Rail to improve performance – the way forward is to rebalance volume and performance incentives.

B.2. Practical Issues

- § The finer details of the charge (eg timing of peak/off-peak periods) could make a big difference.
- § Charging approaches can add significantly to costs in the industry without having a big impact on behaviours. There may be regulatory costs and extra administration costs for Network Rail.
- § The existing capacity charge is crude and has already caused distortions in some areas (eg withdrawing off-peak services).
- § The example CU charges given do not differentiate between different types of traffic on the same line (eg stopping/non-stopping).

B.3. Overall Impact of Charge

- § It is impossible to separate the commercial and non-commercial railway – there has to be a balance. This means that there will always be a role for the government in specifying rail services, which limits the ability of the CU charge to affect TOC behaviour.
- § There are a lot of trade-offs involved in capacity utilisation, for example between performance and the number of paths, and a CU charge could help with this.
- § A CU charge may raise costs for funders and force them to make cuts, probably to off-peak services.
- § A CU charge may only be useful where there is excess demand.
- § Trying to reduce unit costs and improve capacity utilisation may not be the same thing.
- § A CU charge might disadvantage rail in competing against road transport, particularly for freight, as road freight does not face charges based on distance or time of day. Customers have timing requirements, which may require freight to run in peak times. This traffic could be lost to road. In addition, the complexity of the charge makes quote times uncompetitive. The charge would add too much complexity for freight customers.
- § There is a risk of unintended consequences of a CU charge (eg freight rerouting over less suitable lines that leads to an increase in Network Rail's costs).
- § Existing charges and rules have been added incrementally and so this might be an opportunity to revisit the whole approach to incentives in this area.

B.4. Alternatives to CU Charge

- § Specific off-peak problems could be addressed individually rather than through a CU charge
- § The RUS process is a tool for driving decisions. It could be improved rather than introducing a CU charge.
- § Other mechanisms such as revenue sharing could be used instead of a CU charge.
- § A trading approach or buy-back mechanism was suggested.

Appendix C. Detailed Questions Raised by ORR

D.60 Issues to consider in improving incentives to make best use of existing capacity include:

- § whether greater prominence should be given to Network Rail's role as the network system operator by for, example, enhanced licence conditions or a separate licence;
- § whether a single body should be responsible for whole system capacity utilisation, covering both train operations and rail network;
- § whether incentives should be placed on train operators as well as Network Rail to improve capacity utilisation;
- § increasing the power of the volume incentive, for example by exposing Network Rail to downside as well as upside risk;
- § developing a new metric for capacity utilisation, as recommended by the vfm study. Existing measures (such as the capacity utilisation index) are good at identifying where there may be spare capacity in the timetable; however, they are less suitable where a new timetable is required. There may also be benefit in setting explicit targets for the new metric;
- § exposing Network Rail to some form of revenue sharing as part of the cost and revenue sharing proposals, then Network Rail's incentives to accommodate additional traffic would be increased (see below); and
- § amending the structure of charges so that Network Rail's incentives to grow or develop the network are improved. This could encompass: market-based investment incentives by auctioning off capacity, amending variable charges so that they reflect long-run marginal costs and the introduction of scarcity or reservation charges.

Appendix D. Questions Considered at the Stakeholders Workshop

How would train operators' decisions be affected by a CU charge?

Might some passenger operators run fewer, longer trains, or some freight operators switch to less congested periods/routes?

- how widespread would such changes be?
- what might make such changes more/less likely?

Might some passenger operators change their service patterns?

- to focus direct services on main destinations (“hub and spoke” system)?
- to make other changes?

How might these impacts be affected by other regulatory/contractual mechanisms (eg franchise specification, ORR policy on access rights/competition, RUSs, etc)?

How would Network Rail's behaviour be affected by a CU charge?

Would Network Rail be more proactive in identifying additional train paths?

- how widespread would such changes be?
- what might make such changes more/less likely?
- what might discourage this (eg performance risk)?

Would Network Rail start suggesting changes to existing services to free up capacity?

- how widespread would such changes be?
- what might make such changes more/less likely?
- what might discourage this (eg franchise specification, ORR processes)?

How might these impacts be affected by other regulatory/contractual mechanisms (eg volume incentive, licence enforcement/PPM targets, franchise specification, ORR policy on access rights/competition, RUSs, etc)?

What would be the wider impact of exposing Network Rail to stronger financial incentives to maximise the value generated from existing capacity?

What practical issues would need to be addressed?

For the simple options

- can congested parts of the network / times of day be identified with sufficient accuracy?
- is there a risk that some services will be priced off the network unnecessarily?
- how significant are the initial and ongoing implementation costs?

For the more complex option

- can capacity utilisation be measured with sufficient precision?
- is the greater complexity justified by the likely impact of such a charge?

Implementation options

- should the charge be set at a deliberately low level at first?
- how frequently would charges (and charge categories) need to be reviewed?

Overall impact - would a CU charge improve capacity utilisation?

Would it have a sufficiently strong impact (bearing in mind that it would not aim to eliminate all excess demand and the existence of other regulatory/contractual mechanisms)?

Is there any risk of unintended consequences (for example because some services are priced off the network unnecessarily, or because improving capacity utilisation requires operational rather than commercial processes)?

How would a CU charge affect other objectives

- promoting competition between train operators?
- promoting rail freight?
- protecting socially-beneficial services (eg to remote areas)?

Are there other options that we should consider

- something more detailed than Option 2, but not a bespoke or modelling-based approach?
- a different approach altogether (eg a different type of charge, or something based on other parts of the existing charging framework)?

NERA

Economic Consulting

NERA Economic Consulting
15 Stratford Place
London W1C 1BE
United Kingdom
Tel: +44 20 7659 8500
Fax: +44 20 7659 8501
www.nera.com