Response to Network Rail's Periodic Review 2013 – Consultation on the Capacity Charge (July 2012)

By John Haith PhD Student, Institute of Transport Studies, University of Leeds

I am a PhD Student in my Third Year.

My subject is:-

UNDERSTANDING CAPACITY USE AND PERFORMANCE AND THEIR IMPLICATIONS FOR THE PRICING OF CONGESTED RAIL NETWORKS

This involves investigating the effectiveness of different measures of capacity use at predicting timetable reliability and then using the conclusions to suggest the most appropriate form of congestion charge.

I have written some comments about the data set used in my research in Appendix A at the end of this document. Appendix B lists the specific comments in your document that I have used as the basis for this response.

I came across your Consultation Document whilst preparing an abstract for a forthcoming transport conference.

Hopefully, the following comments will be of interest.

Definitions

I refer to a number of different capacity measures in my response. These are:-

- **Original CUI** – The methodology used in previous calibrations of the capacity charge. This uses 'compression' and measures capacity use on relatively short plain-line sections.
- **Original CUI +** - The original methodology with a number of 'improvements' which I believe increase the accuracy of the result.
- **Junction (Station Throat) CUI** – measuring the capacity use at junctions (station throats) using the CUI compression methodology. The calculation is either 'stand-alone' or as part of the calculation of capacity use on adjacent route sections.
- **Heterogeneity (HET)** – the measurement of the minimum gap between successive services.

Précis of my Comments

- I too have found that the Original CUI methodology produces reasonable results.
- I believe that its accuracy can be improved from that obtained previously with some simple changes to the methodology.
- Original CUI has two obvious drawbacks:-
    - i) It excludes junction and station capacity usage.
    - ii) In using 'compression' it eliminates the very thing that determines whether reactionary delay occurs i.e. the gap between services.
- It is my belief that Junction CUI, although technically feasible, will be very difficult to use at a national level in the timescales you have outlined. In any case at this stage I believe that the addition of Junction CUI will give a less accurate result than just using 'Original CUI +'.

- I do not believe that there are any other determinants of reactionary delay. In any case introducing route section specific variables to try and account for observed anomalies is fraught with problems. There needs to be a consistent approach.
- In my research I am testing as many functional forms as possible before deciding which is the most appropriate to describe the relationship between capacity usage and levels of reactionary delay.
- In my Literature Review I identified a number of different metrics for measuring capacity usage. One of them is based on measuring the gaps between services.
- At the moment the heterogeneity methodology (HET) that has been developed, consistently performs best at predicting levels of reactionary delay.
- I believe that HET can be improved further.
- One reason that HET is successful is that it can help explain the apparent anomaly in my data set where an off-peak hour has more un-adjusted reactionary delay than the equivalent peak hour. This is because in this particular off-peak hour there is a great deal of 'bunching' of services.
- The transferability of HET needs to be looked at before it can be considered suitable for wider use.
- There are examples from the current timetable where off-peak 'bunching' can be seen.
- From the point of view of my timescales I would recommend investigating the use of HET to calibrate the Capacity Charge. (I believe that this is the conclusion I will state in my PhD Thesis)
- From the point of view of your timescales (which are much shorter) my recommendation would be to use 'Original CUI +' but with special attention paid to any resulting anomalies and with suitable adjustments made.
- My recommendation would be to avoid Junction / Station Throat CUI for this recalibration exercise.

Use of Original CUI

My starting point was to adopt the CUI methodology used in the original analysis. However, there were a number of elements in the methodology that I felt could be changed to produce a more accurate result. For example, the original methodology used quite a lot of 'averaging' of capacity usage to produce a CUI %. I can understand why this was done. However, in my analysis I have avoided 'averaging' data. This 'CUI +' methodology produces a good result, using the 'success criteria' adopted in the original analysis.

However, there are two obvious drawbacks with the Original CUI methodology:-

1. It is plain-line sections only. As you identify junction and station throat moves are not included in the calculations.
2. It compresses the 'above planning headway' gaps in the timetable. As you identify in Technical Box 1 of your Consultation Document [Page 9] it is the gap between services that determines whether reactionary delay is incurred.

Junction (& Station Throat) CUI

The calculation of CUI for junctions and stations is technically feasible and is the subject of at least one published academic paper (see key references at the end of this document).

I am just beginning work on Junction CUI for my data set.

However, I have already encountered a number of issues:-

1) Junction Capacity Use is much more complicated than Plain Line Capacity Use.
   The margins for Hitchin (At Grade) and Newark Flat Crossing are fairly complex and comprise in each case a number of 'if/then' statements. This makes the calculation of Junction CUI more time consuming than the 'Original CUI +'.

2) There are different types of junctions.
   The two junctions in my data set have substantially different characteristics. For example, Newark Flat Crossing has East – West traffic crossing the ECML whilst Hitchin (At Grade) Junction has Cambridge traffic joining / leaving the ECML.

3) The big question is how to include Junction CUI in the calculated capacity use.

   I am looking at the following Variants:-

   Junction at the end of a route section.
   Junction at the beginning of a route section
   Junction between two route sections.
   Junction as an independent CUI calculation.

   (I have also already calculated 'Expanded CUI' which is based on the methodology used by Network Rail's Engineering Enhancement Department. This calculates the CUI for a much longer section than the original CUI methodology and includes junction capacity use).

   I plan to test the effectiveness of each of the variants. A problem will arise if different variants work best for the two different areas or even for different directions of the same area. Even if one variant works best for both junctions it does not necessarily follow that this will be the case for other types of junction. There is therefore the issue of transferability.

4) High Route Section Capacity Use is likely to mean High Junction Capacity Use etc.

   It seems likely that a high CUI on adjacent route sections will mean a high Junction CUI and a low CUI on adjacent route sections will mean a low Junction CUI.

   If statistical' noise is of the type of a High CUI / Low RD or Low CUI / High RD relationship then the inclusion of junction CUI is likely to make this anomaly worse.

Finally**,** as stated previously I have only just started to calculate junction capacity usage. However, even at this very early stage I can see that the relationships appear to be worse than those obtained from 'Original CUI +'.

<u>Other Determinants of Reactionary Delay</u>

My conclusion is that there is only one determinant of reactionary delay. As described in your consultation document in Technical Box 1, the gap between successive trains will determine how much reactionary delay the second train will incur if the first train is delayed.

During my research I have at various points included explanatory variables other than capacity use in my regression analysis.

For example, one reason for picking the ECML was the presence of two Open Access Operators (Hull Trains and Grand Central) and the variable 'Number of Open Access Paths' seemed an obvious one to include in the analysis. Indeed, this variable was found to be significant. However, further testing found that although 'Open Access Paths' was a significant variable for the Welwyn Area (specifically the Up direction) it was not for the Newark Area. 'Open Access Paths' was therefore discarded as a variable because the use of variables that applied to one area but not another obviously raises issues when the question of transferability is considered.

Two other explanatory variables I have looked at :- 'Number of Freight Paths' and 'Number of Junction Crossing moves' were both found to be insignificant in the regression analyses I have run.

Functional Form

I have adopted the approach used in the original analysis. Namely, I am testing a number of different functional forms to identify which works best. At the moment these are the five functional forms used in the original analysis. Although, I expect an exponential form to give the best result (as established by the original analysis and shown by the regression analysis I have completed to date) I am not making any definite conclusions until all my analysis is complete. I am also bearing in mind the comments made by Symonds, during their review for the ORR of the original methodology, about the exponential and power functions.

Heterogeneity (HET)

One of the papers I discovered during my Literature Review was one describing the results of a Dutch Rail Study (see Vromans et al in the Key References). Its key conclusion was that the reliability of a timetable was determined by how capacity was used rather than how much. The authors proposed the interesting concept of measuring the minimum gap between successive trains using reciprocals (for the simple reason that the smaller the gap the bigger the resulting number).

Taking the example of Low Traffic in Technical Box 1 of the Consultation Document [Page 9]:-

If S1 is 15 minutes in front of S2, then a delay of 10 minutes to it will produce no reactionary delay to S2 (if the ROTP headway is 4 minutes then S1 could incur 11 minutes before S2 was affected).

If S1 is 10 minutes in front of S2 then a delay of 10 minutes would produce 4 minutes of reactionary delay to S2 (assuming a ROTP headway of 4 minutes).

If S1 is 7 minutes in front of S2 then a delay of 10 minutes would produce 7 minutes of reactionary delay to S2 (assuming a ROTP headway of 4 minutes).

If S1 is 4 minutes in front of S2 then a delay of 10 minutes would produce 10 minutes of reactionary delay to S2 (assuming a ROTP headway of 4 minutes).

So:-

i) RD = 0, Gap = 15 minutes, Reciprocal = 1/15 = 0.067
ii) RD = 4, Gap = 10 minutes, Reciprocal = 1/10 = 0.100
iii) RD = 7, Gap = 7 minutes, Reciprocal = 1/7 = 0.143
iv) RD = 10, Gap = 4 minutes, Reciprocal = 1/4 = 0.250

In each of the 4 scenarios the calculated CUI would be the same. However, the measurement of the minimum gaps between trains correctly identifies that trains planned closer together will have a higher risk of reactionary delay than those planned further apart.

I have used the Dutch Rail study paper as the basis for producing a methodology for measuring capacity usage. For example the methodology takes into account the planning headway / junction margin. It is still at a basic standard as there are a number of ways that I feel it could be further improved. However, even at a basic standard it out performs any of the capacity measures based on the volume of capacity used (such as CUI) in the regression analyses I have undertaken.

Obviously, any new methodology needs to be transferrable but the 'bunching' of off-peak services is not restricted to my data set. For example, looking at the current timetable in National Rail Enquiries (search conducted 03/09/12) gives the following departures for my Home Station:-

York to Newcastle between 1000 and 1100

Times	Gap to Previous (Mins)
(0958)	-
1032	34
1034	2
1048	14
1053	5
1058	5
(1132)	

York to London KX between 1000 and 1100

Times	Gap to Previous (Mins)
(0955)	-
1001	6
1025	24
1029	4
1055	26
(1129)	

Obviously, these are Public rather than Working Timetables times but the uneven gaps do suggest that there is a greater risk of reactionary delay than might be suggested by the CUI methodology.

I am hopeful that Network Rail will continue to support my PhD with the supply of data required to test this out.

My Recommendation.

Of course this is based on my research and experience to date.

If there was time, I believe that the most effective methodology (of the ones I have considered) for predicting reactionary delay is the one based on the concept of Timetable Heterogeneity. I believe this would address the statistical noise you refer to in your document and the fact that "the charge does not always fully compensate Network Rail for the increased performance risk associated with accommodating new services".

However, I appreciate that you are working to much shorter timescales than I am. My recommendation would be to use 'Original CUI +' but ensure you have a robust system in place for identifying and understanding any anomalies in the results'. Local adjustments could then be made.

My advice would be to avoid Junction / Station Throat CUI for this calibration. Not because it can't be done – it can. But the sheer added complexity of the calculations required on a National basis means I think it would be difficult to accomplish in your timescales. In any case, the initial work I have done suggest that its inclusion will produce a less accurate result than 'Original CUI +'.

I am of course happy to answer any questions about anything contained in this response or about any other aspect of my PhD research.

Regards

John Haith
PhD Student
Institute of Transport Studies
University of Leeds

(September 2012)

Key References used in Writing This Response

Armstrong, J., S.Blainey, J. Preston and I. Hood (2011*) Developing a CUI Based approach to Network Capacity Assessment* – 4th International Seminar on Railway Operations Modelling and Analysis (Rome 2011)

Faber Maunsell (2007) Capacity Charge Tariff PR2008 - Recalculating the Capacity Charge Tariff for PR2008

Gibson, S., G.Cooper and B.Ball (2002) *The evolution of capacity charges on the UK rail network* Journal of Transport Economics and Policy Vol.36 Part 2 May 2002 pp 341-354.

Krueger, H., (1999) "Parametric Modelling in Rail Capacity Planning" Proceedings of the 1999 Winter Simulation Conference pp1194-1200

Network Rail (2012) Periodic Review 2013 – Consultation on the Capacity Charge (Downloaded from the Internet 24/08/12)

ORR (2000) The Periodic Review of Railtrack's access Charges – Capacity Charges – A technical consultation document.

Vromans, M.J.C.M., R.Dekker, and L.G.Kroon, (2006) *Reliability and heterogeneity of railway services* European Journal of Operational Research 172 2006 pp 647-655

APPENDIX A

My Data Set

For my research I am looking at two sections of the East Coast Main Line. One 46 mile section is centred on Newark Flat Crossing. The other 31 mile section is centred on Hitchin (At Grade) Junction and Welwyn Viaduct. The resulting data set is small but complex. The reason for using a small data set is that the large number of regression runs required is relatively manageable. Secondly, it has allowed me to really get to know the data. This is particularly important when attempting to understand any anomalies. The reason for using a complex data set is on the basis that any conclusions are more likely to be transferrable to other areas than if a simple data set was used.

One aspect I am particularly interested in is the appropriate size of route section. To examine this I am calculating capacity usage at both an area level (i.e. the ECML sections in their entirety) and a route section level (i.e. the ECML sections subdivided into the equivalent of Original CUI route sections).

I am using December 2008 SX Timetable data.

I am using hourly time periods. As noted by Gibson, Cooper and Ball in their paper this is the smallest practical time period for capacity measurement. This gives the most accurate picture of the size and duration of peak and off-peak periods.

All Timetable Data and Reactionary Delay data used in the analysis was kindly supplied by Network Rail.

Using the data set I have already tested numerous variants of the capacity measures identified during my initial Literature Research. Although, there is still much analysis to do I have done sufficient to be able to reach conclusions about the most effective capacity measure for predicting the reliability of a timetable.

Apparent Anomalies

There are a number of apparent anomalies in my Reactionary Delay data which have a significant impact on the results of my analysis. The biggest apparent anomaly is the fact that the time period with the highest level of observed reactionary delay is an off-peak hour. This means that any capacity measure which calculates the volume of capacity used will struggle to accurately predict the level of reactionary delay.

HET helps explain this apparent anomaly. Services in the off-peak hour with the greatest amount of observed reactionary delay are ‘bunched’ together (9 of the 13 services are the minimum headway behind the train in front). The HET calculation gives a much higher usage figure than if these off-peak services had been evenly spread.

Another interesting anomaly is a high capacity usage in an evening off-peak period in the Welwyn Up area. Although, there is a high capacity usage there is relatively low reactionary delay. The interesting aspect about this time period is that there are several planned FCC ecs moves (these additional trains help explain the relatively high capacity usage). It appears though that reactionary delay was minimised on the main line by ‘holding’ these ecs moves on the Cambridge Branch. None of the capacity usage methodology measures I am currently using account for this. This apparent anomaly does however raise interesting questions about the size and scope of route sections.

APPENDIX B

Key points from your document that my comments refer to

The issues in your consultation document that my comments refer to are as follows:-

- Your proposal to use CUI (because your "testing" indicates that it gives reasonable results and it is consistent with the previous approach). [Page 3 / 19]
- Your comments about 'other determinants of reactionary delay' and your reference to Junction and station throat moves [Page 20]
- Your description of reactionary delay under different traffic volumes in Technical Box 1 [Page 9]
- Functional Form [Page 20]
- Your statement that there are " two types of metric that are typically used to measure capacity" [Page 19]
- The concern that "the charge does not always fully compensate Network Rail for the increased performance risk associated with accommodating new services". [Page 4]
- "Past regression analyses have been characterised by a large amount of statistical noise". [Page 20]
- Your comment that "In practically any transport context ... the notion that higher capacity utilisation gives rise to more reactionary delay is virtually beyond question". [Page 21]