

Pedro Abrantes
Head of Analysis and Rail Economics

22 November 2018

Caitlin Scarlett,
Schedule 8 Recalibration Lead
Rail Delivery Group

Dear Caitlin,

Schedule 8 Recalibration: Approval of CP6 Schedule 8 parameters in the passenger operator regime

1. In your letter of 8 November 2018, RDG requested approval of passenger operator payment rates, passenger operator benchmarks, Network Rail benchmarks and Sustained Poor Performance (SPP) thresholds for use in the passenger operator Schedule 8 regime in CP6.
2. We note that, these parameters have been recalibrated separately for Govia Thameslink Railway (GTR) and London South Eastern Railway (LSER). In the remainder of this letter we refer to the recalibration of the above parameters for all other service groups as the 'national recalibration'.
3. Although the parameters listed above were recalibrated separately for GTR and LSER the methodologies used were broadly similar to the methodologies used in the national recalibration. In addition, parts of the recalibration of these parameters for GTR and LSER were calculated as part of the national recalibration, such as the SPP thresholds.
4. In some cases, operators and Network Rail routes could not reach agreement on the recalibration of the passenger operator and/or Network Rail benchmarks. We discuss these below.
5. We received another submission on 15 November 2018 from the Schedule 8 recalibration lead for GTR and LSER requesting approval of Monitoring Point Weightings and Cancellation Minutes for GTR and LSER service groups.

Passenger operator payment rates

6. We note that, through the passenger operator Recalibration Working Group, operators and Network Rail have had the opportunity to review and challenge the recalibration of the passenger operator payment rates. In particular, we note that no party has objected to the proposed recalibration of the passenger operator payment rates.

7. In addition, we note that the recalibration of the passenger operator payment rates has been independently audited by Vivacity and that they have confirmed that they have no concerns that need to be addressed.
8. Thus, in light of the above and having reviewed both the methodology and audit reports, we approve the recalibration of the passenger operator payment rates for use in CP6.

Passenger operator benchmarks

9. We note that, through the passenger operator Recalibration Working Group, operators and Network Rail have had the opportunity to review and challenge the recalibration of the passenger operator benchmarks. In particular, we note that, with the exception of cases where an operator and Network Rail do not agree on the recalibration of the passenger operator benchmarks, no party has objected to the proposed recalibration of the passenger operator benchmarks.
10. In addition, we note that the recalibration of the passenger operator benchmarks has been independently audited by Vivacity and that they have confirmed that they have no concerns that need to be addressed.
11. Thus, in light of the above and having reviewed both the methodology and audit reports, we approve the recalibration of the passenger operator benchmarks for use in CP6 (subject to what we say in relation to the operators set out below).
12. Below we set out our decisions on the passenger operator benchmarks that operators and Network Rail disagreed on. It should be noted that we received submissions for the disputes discussed below two weeks before the date when a decision had to be reached; the limited time we have had to make a decision has been a factor in the outcomes.

Great Western Railway

13. Great Western Railway (GWR) disagreed with the passenger operator benchmarks calculated for its service groups as part of the national recalibration. The Network Rail Western route agreed with the passenger operator benchmarks recalibrated for GWR's service groups as part of the national recalibration.
14. We note GWR's concern that using unadjusted historical data from the standard recalibration period in the national recalibration (2015-16 and 2016-17) does not reflect its expected performance in CP6.

15. In its submission, GWR suggested an alternative methodology for calculating the passenger operator benchmarks for its service groups to better reflect expected performance in CP6. However, the alternative methodology had not been sufficiently developed and no alternative passenger operator benchmarks had been calculated for GWR's service groups.
16. As a result, we have not been able to determine if GWR's alternative methodology would provide passenger operator benchmarks that, relative to the benchmarks calculated as part of the national recalibration, are a better reflection of expected performance of GWR's service group in CP6.
17. As GWR noted in its submission, there is not sufficient time to further develop its alternative methodology and calculate alternative benchmarks for GWR's service groups.
18. Based on all the considerations above, we have decided that the passenger operator benchmarks for GWR's service groups in CP6 should be the passenger operator benchmarks calculated as part of the national recalibration.
19. However, we will consider GWR's concerns on setting passenger operator benchmarks on the basis of unadjusted past performance and setting a single benchmark for each service group for the duration of the control period during the next recalibration of the passenger operator benchmarks.

Hull Trains

20. Hull Trains disagreed with the passenger operator benchmarks calculated for its service groups as part of the national recalibration. The Network Rail London North Eastern (LNE) route agreed with the passenger operator benchmarks recalibrated for Hull Trains' service groups as part of the national recalibration.
21. We note Hull Trains' concern that the standard recalibration period used as the basis for the recalibration of the passenger operator benchmarks in the national recalibration does not reflect its expected performance in CP6, particularly at the start of the control period.
22. In its submission, Hull Trains proposed an alternative recalibration period as the basis for the calculation of the passenger operator benchmarks for its service groups. However, its submission does not include alternative passenger operator benchmarks.

23. The evidence Hull Trains have provided in its submission demonstrates that its performance has worsened since the standard recalibration period used in the national recalibration. However, based on Hull Trains' submission we are not able to take a view on whether Hull Trains' performance will continue at this level in CP6, or return to the level during the recalibration period.
24. In addition, the standard recalibration period used in the national recalibration was selected by the passenger operator Recalibration Working Group in July 2017. In addition, operators had opportunities in Summer 2017 and again in Spring 2018 to select a bespoke recalibration period if they demonstrated that the recalibration period of 2015-16 and 2016-17 was not a good indication of expected performance in CP6.
25. Moreover, due to the late stage we are at in the recalibration process there is not sufficient time to calculate alternative benchmarks for Hull Trains' service groups using its proposed alternative recalibration period. If we did have time, Hull Trains would still need to demonstrate to us that any alternative passenger operator benchmarks better reflect its expected performance in CP6.
26. Based on all the considerations above, we have decided that the passenger operator benchmarks for Hull Trains' service groups in CP6 should be the passenger operator benchmarks calculated as part of the national recalibration.
27. We note that in its submission Hull Trains proposed a mid-control period recalibration in CP6 once its new fleet is in place. When its new fleet comes into service in CP6 Hull Trains is able to propose a Schedule 8 recalibration through existing contractual provisions, specifically paragraph 17 of Schedule 8 of passenger track access contracts. More information on 're-opening' Schedule 8 within a control period is available in our PR18 final determination overview document (available [here](#)).

Northern

28. Northern disagreed with the passenger operator benchmarks calculated for its service groups as part of the national recalibration. The Network Rail LNE route agreed with the passenger operator benchmarks recalibrated for Northern's service groups using the national recalibration.
29. We note Northern's concern that the significant changes to its services since the standard recalibration period used in the national recalibration means its benchmarks will not reflect its expected performance in CP6.

30. In its submission, Northern proposed using an alternative recalibration period as the basis for the calculation of the passenger operator benchmarks for its service groups. However, the calculation of passenger operator benchmarks using the alternative recalibration period had not been undertaken at the time of Northern's submission.
31. As discussed in paragraph 24, the standard recalibration period used in the national recalibration was selected by the passenger operator Recalibration Working Group and operators had opportunities to select a bespoke recalibration period if they could demonstrate that the standard recalibration period was not a good indication of expected performance in CP6.
32. Due to the late stage we are at in the recalibration process there is now not sufficient time to calculate alternative benchmarks for Northern's service groups using its proposed alternative recalibration period. If we did have time, Northern would still need to demonstrate to us that any alternative passenger operator benchmarks better reflect its expected performance in CP6.
33. Based on all the considerations above, we have decided that the passenger operator benchmarks for Northern's service groups in CP6 should be the passenger operator benchmarks calculated as part of the national recalibration.

Network Rail benchmarks

34. We note that, through the passenger operator Recalibration Working Group, operators and Network Rail have had the opportunity to review and challenge the recalibration of the Network Rail benchmarks. In particular, we note that, with the exception of cases where an operator and Network Rail do not agree on the recalibration of the Network Rail benchmarks, no party has objected to the proposed recalibration of the Network Rail benchmarks.
35. In addition, we note that the recalibration of the Network Rail benchmarks has been independently audited by Vivacity and that they have confirmed that they have no concerns that need to be addressed.
36. Thus, in light of the above and having reviewed both the methodology and audit reports, we approve the recalibration of the Network Rail benchmarks for use in CP6 (subject to what we say in relation to the operators set out below).
37. Below we set out our decisions on the Network Rail benchmarks that operators and Network Rail disagreed on. It should be noted that we received submissions for the disputes discussed below two weeks before the date when a decision had

to be reached; the limited time we have had to make a decision has been a factor in the outcomes.

Chiltern Railway

38. Chiltern Railway proposed an alternative Network Rail benchmark for service group HO04. The Network Rail London North Western (LNW) route agreed with the Network Rail benchmark calculated for service group HO04 as part of the national recalibration.
39. In the national recalibration, a linear regression was used to evaluate the relationship between actual minutes lateness (AML)¹ and service group-level Network Rail delay per 100 train kilometres, where the adjusted R-squared value for the service group was 0.70 or higher. For service groups that had an R-squared value for the regression below 0.70 Steer presented options to the operator and the Network Rail route. If no agreement was reached the regression was not used in the calculation to apply Network Rail's consistent route measure of passenger performance (CRM-P) CP6 baseline trajectories to the Network Rail benchmarks.
40. For service group HO04 Chiltern Railway proposed to use the regression approach, despite the R-squared for the regression for this service group being slightly below the 0.70 threshold (0.67).
41. Although the R-squared value for this service group is only slightly below the 0.70 threshold Chiltern Railway have not provided evidence on why it is still appropriate to use a the regression for this service group. As a result, we are not able to accept Chiltern Railway's proposal.
42. Based on all the considerations above, we have decided that the Network Rail benchmark for service group HO04 in CP6 should be the Network Rail benchmark calculated without using the regression.

c2c

43. c2c and the Network Rail Anglia route proposed alternative Network Rail benchmarks for c2c's HT01 peak service group.
44. The adjusted R-squared value for the regression used to convert service group-level Network Rail delay per 100 kilometres trajectories to AML for this service group was below the 0.70 threshold set for using the regression approach (0.54).

¹ Steer used the term Performance Minutes (PM) to describe the sum of AML and DML rather than the often-used term Average Minutes Lateness, so in this context AML refers to actual minutes lateness.

Steer provided c2c and the Network Rail route with three options: use the default approach of not using the regression; use the regression with c2c's bespoke recalibration period (2016-17 only); or use the regression using the recalibration period of 2016-17 period 10 to 2017-18 period 9.

45. c2c proposed using the "with regression" approach with the recalibration period of 2016-17 period 10 to 2017-18 period 9, while the Network Rail Anglia route proposed using the "with regression" approach with the c2c's bespoke recalibration period.
46. The R-squared value for the approach with the regression proposed by c2c is above the 0.70 threshold (0.88). However, we have not been provided evidence on why it is appropriate to use 2016-17 period 10 to 2017-18 period 9 as the recalibration period.
47. The Network Rail Anglia route's rationale for its proposal is that it considers that there is no reason to use a different recalibration period for this regression. We do not consider this as evidence either in support of using its proposed "with regression" approach with the c2c's bespoke recalibration period, or against using the default "without regression" approach.
48. As a result, we are not able to accept c2c's or the Network Rail Anglia route's proposed Network Rail benchmark for c2c's HT01 peak service group.
49. Based on all the considerations above, we have decided that the Network Rail benchmark for HT01 peak in CP6 should be the Network Rail benchmark calculated without using the regression (i.e. the Network Rail benchmark calculated using the national default approach).

GTR

50. GTR disagreed with the Network Rail benchmarks calculated for its service groups as part of the national recalibration. The Network Rail South East route agreed with the Network Rail benchmark calculated for GTR's service groups as part of the national recalibration.
51. We note that GTR was concerned with the Network Rail CRM-P CP6 baseline trajectories. In particular, it was concerned that the GTR performance plan is mainly based on the Network Rail South East route with limited input from other routes it operates on. In addition, due to the diverse range of services that GTR operates, GTR considered it inappropriate to set a single Network Rail trajectory

- for its services. GTR requested additional time to change the Network Rail CRM-P CP6 baseline trajectory for its services.
52. We have already determined the Network Rail CRM-P CP6 baseline trajectories for each route as part of our PR18 final determination. These will act as a baseline against which we will measure Network Rail's delivery to current and future passengers and freight end users over the control period in our monitoring and reporting. Although these baselines are applied to the Network Rail benchmarks, it is important to stress that the level of the trajectories are not within the scope of the Schedule 8 recalibration. When setting Network Rail's CRM-P CP6 baseline trajectories, prior to our final determination, we separately took into account evidence provided by operators in response to our PR18 draft determination.
 53. GTR also raised a concern about the recalibration period used as the basis for the Network Rail benchmarks for its service groups. GTR did not consider it appropriate to use 2015-16 as part of the recalibration period, due to the significant changes to its services since then.
 54. We recognise there have been changes to GTR's services since the recalibration period. However, as explained in paragraph 24, this period was selected by the passenger operator Recalibration Working Group and operators have had opportunities to demonstrate why they should have a bespoke recalibration period.
 55. GTR also explained that service group re-mappings and other changes to its services in CP6 will mean the Network Rail benchmarks calculated for CP6 the national recalibration will deviate significantly from its actual performance in CP6.
 56. We recognise there will be significant changes to GTR's services in CP6. In our PR18 final determination we explained that we would be minded to approve applications for a large scale mid-control period recalibration of Schedule 8 for significant changes in traffic on the network, for example, as a result of the addition of new GTR services.
 57. In its submission, GTR have not proposed alternative Network Rail benchmarks for its service groups. Due to the stage we are at in the recalibration process we are not able to provide GTR with additional time to develop a methodology and calculate alternative Network Rail benchmarks for its service groups. If we did have time GTR would still need to demonstrate to us that any alternative Network Rail benchmarks better reflect Network Rail's expected performance to its service groups in CP6.

58. Based on all the considerations above, we have decided that the Network Rail benchmark for GTR's service groups in CP6 should be the Network Rail benchmarks calculated as part of the national recalibration.

Northern

59. Northern disagreed with the Network Rail benchmarks calculated for its service groups as part of the national recalibration. The Network Rail LNE route agreed with the Network Rail benchmark calculated for Northern's service groups as part of the national recalibration.
60. Northern's main concern about the Network Rail benchmarks relates to Network Rail's CRM-P CP6 baseline trajectories. As discussed above, we have already set Network Rail's CRM-P CP6 baseline trajectories as part of the PR18 final determination.
61. In addition, we took the arguments Northern provided in its submission on the Network Rail benchmarks into account when we set Network Rail's CRM-P CP6 baseline trajectories as part of our final determination.
62. In its submission, Northern have not proposed alternative Network Rail benchmarks for its service groups. Due to the stage we are at in the recalibration process we are not able to provide Northern with additional time to develop a methodology and calculate alternative Network Rail benchmarks for its service groups. If we did have time Northern would still need to demonstrate to us that any alternative Network Rail benchmarks better reflect Network Rail's expected performance to its service groups in CP6.
63. Based on all the considerations above, we have decided that the Network Rail benchmark for Northern's service groups in CP6 should be the Network Rail benchmarks calculated as part of the national recalibration.

Sustained Poor Performance (SPP) thresholds

64. We note that, through the passenger operator Recalibration Working Group, operators and Network Rail have had the opportunity to review and challenge the recalibration of the SPP thresholds. In particular, we note that, with the exception of GWR, no party has objected to the proposed recalibration of the SPP thresholds. We discuss GWR's submission separately in paragraphs 67-71 below.

65. In addition, we note that the recalibration of the SPP thresholds has been independently audited by Vivacity and that they have confirmed that they have no concerns that need to be addressed.
66. Thus, in light of the above and having reviewed both the methodology and audit reports, we approve the recalibration of the SPP thresholds for use in CP6.

GWR

67. On 31 October 2018, GWR submitted a proposal for an alternative methodology for calculating the SPP thresholds for the first year of the control period.
68. Passenger operators can make an SPP claim when the average Schedule 8 payment made by Network Rail over 13 periods exceeds the amount which would be paid if performance were a certain proportion (20% for CP6) worse than benchmark. As the SPP thresholds are based on Network Rail's performance relative to its benchmarks over 13 periods any claims in the first year of the control period also takes into account Network Rail's performance in the last year of the previous control period.
69. GWR's rationale for its proposal was that the methodology proposed by the passenger operator Recalibration Working Group did not factor in the Network Rail payment rates and benchmarks from the last year of CP5. As a result, under the Working Group's proposal the SPP thresholds in the first year of CP6 are only based on Network Rail's payment rates and benchmarks in the first year of CP6.
70. We recognised there was merit in GWR's proposal and several passenger operators supported this proposal. However, a number of members of the passenger operator Recalibration Working Group raised concerns that they had not had sufficient time to review it. As a result, we were not able to accept GWR's proposal. As a result, we decided the SPP thresholds in the first year of GWR should be calculated using the same methodology as the national recalibration.
71. It should be noted that we will revisit this proposal during a future recalibration of the SPP thresholds.

Monitoring Point Weightings and Cancellation Minutes for GTR and LSER service groups

72. We note that GTR, LSER and Network Rail have had the opportunity to review and challenge the recalibration of the Monitoring Point Weightings and Cancellation Minutes for GTR and LSER service groups. In particular, we note



that no party has objected to the proposed recalibration of the Monitoring Point Weightings and Cancellation Minutes.

73. In addition, we note that the recalibration of the GTR and LSER Monitoring Point Weightings and Cancellation Minutes has been independently audited by Vivacity and that they have confirmed that they have no concerns that need to be addressed.
74. Thus, in light of the above and having reviewed both the methodology and audit reports, we approve the recalibration of the GTR and LSER Monitoring Point Weightings and Cancellation Minutes.

Yours sincerely,

Pedro Abrantes