Deren Olgun
Principal Economist

4 May 2018

Caitlin Scarlett
Schedule 8 Recalibration Lead
Rail Delivery Group

Dear Caitlin,

Final decision on the Delay Multipliers to be used in the recalibration of the Schedule 8 Network Rail payment rates in the passenger operator regime

1. In your letter of 2 February 2018 you asked us to determine (for all applicable flows) which set of Delay Multipliers should be used for the calculation of the Marginal Revenue Effects (MREs), as part of the recalibration of the Schedule 8 Network Rail payment rates. Your letter records the process by which Network Rail and operators agreed to ask ORR to determine this issue, including the delegation of these decisions to the 'revenue sub-group' of the Recalibration Working Group.

2. You put forward the following options for our consideration:

    a. Use the Delay Multipliers as set out in PDFH v5.1.

    b. Use the Delay Multipliers that were originally issued alongside PDFH v6.0.[1]

    c. Use an unweighted average of the Delay Multipliers in (a) and (b), above.

3. We have considered the arguments in the annexes to your letter for and against the different options.

4. Whilst we recognise that there may be weaknesses in the PDFH v5.1 Delay Multiplier evidence base we are, on the basis of the arguments put to us, of the view that the alternative options (i.e. (b) and (c)) are more likely to generate inaccurate estimates of the 'true' Delay Multipliers than option (a). For this reason, we have concluded that, of the three options above, the Delay Multipliers described in option (a) should be used as the basis for the recalibration.

5. As requested, we informed you of this decision by email on 16 February 2018, with a brief explanation of our rationale, to meet your deadline. This letter provides further information on our rationale as well as confirming this decision.

---

[1] We note that these are not published in the final version of PDFH v6.0, owing to a change in methodology that obviates the need for Delay Multipliers.

Background

6.   Delay Multipliers reflect the relative values that passengers on particular service types place on planned versus unplanned changes to journey time. For instance, a Delay Multiplier of 2 for a particular sort of service indicates that passengers on services of that sort are indifferent between two minutes of planned additional journey time or one minute of unplanned additional journey time (i.e. one minute of lateness).

7.   In CP5, Delay Multipliers were used in the calculation of the MREs of a minute's lateness for each flow, which, in turn, were used to calculate the Schedule 8 Network Rail payment rate for each service group.

8.   For CP6, you have proposed that this methodology should be used for all flows except London and South East commuter flows, which will be calculated using a different methodology that does not require the use of Delay Multipliers.

9.   The Delay Multipliers used in CP5 were those set out in PDFH v5.1.

10.   We understand that a set of revised Delay Multipliers were originally issued alongside PDFH v6.0 but were removed when the adoption of AML elasticities rendered them unnecessary. As you note in your letter: 'these Delay Multipliers are the unweighted average of the PDFH v5.1 Delay Multipliers, and a set of Delay Multipliers which takes evidence from the same studies as PDFH v5.1, but enters an elasticity of 0 for each study which produced a statistically insignificant result.'

11.   Our understanding is as follows: There are two sets of evidence on which the different estimates of the Delay Multipliers are based. The first set excludes the results of studies that did not find a statistically significant relationship between performance (however it was measured) and revenue. We will hereafter refer to the estimates of the Delay Multipliers based on this evidence as the *exclusive* estimates.

12.   The second set of evidence includes all the evidence in the first set, as well as studies that did not find a statistically significant relationship, for which an elasticity (between performance and revenue) of zero is recorded. We will hereafter refer to the estimates of the Delay Multipliers based on this second set of evidence as the *inclusive* estimates.

13.   The Delay Multipliers referred to in PDFH v5.1 are the exclusive estimates. The Delay Multipliers that your letter describes as 'the Delay Multipliers which were

originally issued alongside PDFH v6.0' are an unweighted average of the exclusive and inclusive estimates.

14.   We understand that both sets of evidence were available at the time of PDFH v5.1, and were considered by the team that compiled it.

Summary of the dispute

15.   Your letter noted that operators and Network Rail could not reach agreement on which set of Delay Multipliers should be used for the recalibration of the Schedule 8 Network Rail payment rates for the applicable flows. As a result, you asked us to determine this issue in accordance with the process industry has agreed for resolving such disputes.

16.   You put forward the following options for our consideration:

a.   Use the Delay Multipliers as set out in PDFH v5.1.

b.   Use the Delay Multipliers that were originally issued alongside PDFH v6.0.

c.   Use an unweighted average of the Delay Multipliers in (a) and (b), above.

17.   Characterising these options in the terms used above:

a.   Use the exclusive estimates.

b.   Use an unweighted average of the inclusive and exclusive estimates.

c.   Use a *weighted* average of the inclusive and exclusive estimates, with twice as much weight placed on the exclusive estimates as on the inclusive estimates.[2]

18.   Operators were in favour of option (a), whilst Network Rail was in favour of option (b). Since we received no proposals supporting option (c) we have not considered it further, although many of the remarks that follow are also relevant to that option.

19.   We have reviewed the arguments put forward by both operators and Network Rail in the annexes to your letter, and, on the balance of those arguments, we have determined that the Delay Multipliers described in option (a) should be used.

---

[2] The Delay Multipliers in option (c) take the arithmetic mean of option (a) and (b) delay multipliers, which are, respectively, the exclusive estimates and the arithmetic mean of the inclusive and the exclusive estimates. Thus, the Delay Multipliers in option (c) are a *weighted* average of the inclusive and exclusive estimates, with twice the weight placed on the exclusive estimates.

Network Rail's arguments

20.   In Annex 2 to your letter, Network Rail argued that option (b) should be adopted, on the following basis:

   i.   The exclusive estimates likely overstate the 'true' Delay Multipliers and should be treated as an 'upper bound';

   ii. The inclusive estimates represent a 'lower bound';

   iii. The arithmetic mean of the inclusive and exclusive estimates, therefore, provides the best approximation of the 'true' Delay Multipliers.

Our response

21.   With respect to (i), while we recognise that there is some evidence to support the claim that the exclusive estimates overstate the 'true' Delay Multipliers; the evidence is far from definitive. This claim is, as the revenue subgroup agreed, and as the operators stressed in their Annex 1 to your letter, supported by a very small number of studies; we consider that more evidence would be required to demonstrate it reliably.

22.   We note that Network Rail also sought to support this claim by arguing that the exclusion of studies that did not find a statistically significant relationship from the evidence base may lead to inflated values, because insignificant elasticities will tend to be lower. We are not convinced that this is the case. As we noted in the meeting of the revenue sub-group on 25 January 2018 (hereafter: the revenue sub-group meeting), it is not clear to us why insignificant elasticities should tend to be lower, given that statistical significance (or the lack thereof) is also a function of the size of the relevant standard errors. Further, in Annex 1 to your letter, Network Rail did not demonstrate, to our satisfaction, why insignificant elasticities should tend to be lower.

23.   For these reasons we are not persuaded that the exclusive estimates overstate the 'true' Delay Multipliers (i.e. we are not persuaded that (i) is true), though we recognise it as a possibility.

24.   With respect to (ii), we do not agree that the inclusive estimates represent a reliable or credible 'lower bound'.[3]

---

[3] In the absence of a clear definition, we assume that the 'lower bound' of the Delay Multipliers refers to the lowest value that the 'true' Delay Multipliers could plausibly be.

25. First, as we noted in the revenue sub-group meeting, we are very concerned about the decision to use zeroes in the place of insignificant estimates, as was done in the inclusive estimates. Network Rail asserted that: 'we consider that it is entirely appropriate in most cases to assume an elasticity of 0 for a statistically insignificant result – to use anything but 0 would mean that we are artificially imposing a relationship, when the evidence suggests there is none.' This is incorrect – the evidence referred to does not suggest that there is no relationship; it suggests only that we cannot rule out, at a given degree of confidence, the possibility that there is no relationship. Moreover, as the operators note in their Annex 1 to your letter: 'a choice of zero is completely arbitrary: even if zero falls within the confidence interval, any other number within the same interval could have been chosen with no less justification.' Finally, the interpretation of a Delay Multiplier of zero is that passengers would be willing to tolerate any amount of delay in order to avoid a minor increase in their planned journey time. This is plainly absurd, and it is thus hard to see how the use of zeroes for insignificant results could ever generate reliable estimates. For these reasons, we do not find the inclusive estimates to be reliable estimates of even a lower bound for the Delay Multipliers.

26. Furthermore, we do not think the inclusive estimates represent credible estimates of a lower bound for the Delay Multipliers. As discussed in the revenue sub-group meeting, some of the inclusive estimates of the Delay Multipliers are less than 1. This implies that passengers prefer unplanned changes in journey time to planned changes in journey time of equal magnitude. This is extremely implausible. We suggest that a credible 'lower bound' is set at the lowest value that the 'true' Delay Multipliers could plausibly be. As operators also note, the lowest plausible value of any of the Delay Multipliers is 1 (meaning that passengers are indifferent between planned and unplanned increases in journey time); so the fact that some of the inclusive estimates are below 1 makes them too low to be a credible lower bound.

27. Finally, with respect to (iii), even if we were satisfied that the exclusive and inclusive estimates do represent upper and lower bounds, respectively, it does not follow that the arithmetic mean of the two is the best estimate of the 'true' Delay Multipliers, or even that it is a better estimate of the 'true' Delay Multipliers than the exclusive estimate.

28. The arithmetic mean assumes that the 'true' Delay Multipliers are equidistant from the upper and lower bounds. However, this assumption is not warranted by any of Network Rail's arguments; there may be good reasons for thinking that the true values are closer to the upper bound, in which case the use of the arithmetic mean

would not yield the best estimate. For this reason, even if we were persuaded that it is appropriate to treat the exclusive and inclusive estimates as upper and lower bounds, respectively, we are not persuaded that it follows that the arithmetic mean of the estimates will produce a more accurate estimate of the 'true' Delay Multipliers than taking the exclusive estimates on their own.

Our decision

29. We have considered the views on the relative merits of option (a) and (b) set out in the Annexes to your letter. For each and all of the reasons above, we have concluded that option (b) is more likely (or, at least, no less likely) to generate inaccurate estimates of the 'true' Delay Multipliers than option (a).[4] For that reason, we have determined that the Delay Multipliers described in option (a) should be used for the recalibration of the Schedule 8 Network Rail payment rates for all applicable flows.

Next steps

30. This letter states our final decision. This decision is restricted only to the question of which set of delay multipliers (of the options proposed) should be used. You should note that we will still need to review and approve the detail of how the payment rates (and associated elements of the regime) have been calculated once they have been calculated (as per the approach to the recalibration that we have previously discussed).

31. Finally, I would like to thank you for leading the recalibration of the regime and bringing this dispute to our attention. I would like also to thank all involved for respecting the agreed process for resolving disputes.

Yours sincerely,

Deren Olgun

---

[4] We consider that, for similar reasons, option (c) is also more likely to generate inaccurate estimates of the 'true' Delay Multipliers than option (a). As noted, we have not discussed this since we received no proposals in support of option (c).