

RDG Sch8 Audit - NRPR report

1.0 JUNE 2018



vivacity RAIL CONSULTING

RDG SCH8 AUDIT - NRPR
REPORT

0	Executive Summary	1
1	Introduction	3
1.1	BACKGROUND	3
1.2	THE AUDIT - TERMS OF REFERENCE	3
1.3	THIS REPORT	4
2	Audit Approach	5
2.1	VRC'S OVERALL AUDIT APPROACH	5
2.2	NRPR MODELS AND PROCESS - APPROACH TO THE AUDIT	5
2.2.1	INTRODUCTION	5
2.2.2	OVERVIEW OF THE SDG PROCESS AND MODEL SUITE	6
3	Audit Review and Findings	10
3.1	INTRODUCTION	10
3.2	AUDIT OF THE DATA PREPARATION PIPELINE	10
3.2.1	WORK DONE	10
3.2.2	OBSERVATIONS	10
3.3	AUDIT OF THE CONTROLLING PROCESS	12
3.3.1	WORK DONE	12
3.3.2	OBSERVATIONS	12
3.4	AUDIT OF THE INPUT DATA TO THE NRPR CALCULATION MODEL	13
3.4.1	WORK DONE	13
3.4.2	OBSERVATIONS	14
3.5	AUDIT OF THE MODEL METHODOLOGY	15
3.5.1	WORK DONE	15
3.5.2	OBSERVATIONS	15
3.6	DETAILED AUDIT OF MODEL FUNCTIONALITY VERSUS EXPECTED RESULTS	16
3.6.1	WORK DONE	16
3.6.2	OBSERVATIONS	17
3.7	SUMMARY OF ISSUES RAISED	18
3.8	CONCLUSIONS AND RECOMMENDATIONS	19
3.8.1	CONCLUSIONS - NRPR CALCULATION	19
3.8.2	RECOMMENDATIONS	20
	Appendix A Items Reviewed	21
A.1	SOURCE DOCUMENTATION AND METHODOLOGY	21
A.2	DATA PIPELINE ELEMENTS	22
A.3	TOC-SPECIFIC DATA PREPARATION AND MODEL CONTROL	23
A.4	NRPR MODEL	24
	Appendix B Methodology Checks	26
	Appendix C Data pipeline log	34
C.1	REVENUE TOTALS BY STAGE AND OVERALL MISMATCH	35
C.2	MISMATCHES BY STAGE	36

C.3	RESIDUAL FLOWS	37
Appendix D Decision Tables and Test Cases		39
D.1	INTRODUCTION	39
D.2	DECISION TABLES	39
D.3	TEST CASES - FLOW DATA	41
D.3.1	DECISION TABLE CASES	41
D.3.2	OTHER TEST CASES	42
D.4	THE TEST HARNESS	42
Appendix E Results of Detailed Tests		43

RDG SCH8 AUDIT - NRPR
REPORT

VERSION HISTORY

Version	Date	By	Notes
0.1	29/05/2018	PJ	First draft
0.2	01/06/2018	PJ	Revised following SDG and RDG comments and new information on CQs
1.0	07/06/2018	PJ	Release following further SDG comments

0 Executive Summary

Vivacity Rail Consulting (VRC) have audited the process and models used by SDG to calculate Network Rail Payment Rates.

The review has covered the methodology used, the data preparation pipeline, the reference data used in the spreadsheet models, the formulae used in the models and the processes used to create an audit trail to verify correctness of the process.

During the audit we raised 16 issues with RDG. Of these, 5 remain open. 3 of these 5 are Clarification Questions requesting more information. There is 1 issues of severity 3 - material impact on one or more operators, and one of severity 4. Only 1 of these items requires further work by SDG.

The conclusions of our audit are:

- The R code and spreadsheet models have generally been built to a high standard
- The methodology used is sound and as agreed with the Schedule 8 Working Group
- No major issues have been found with the input data, model and calculation
- There were minor areas of concern to do with the location mapping and service group lookup which resulted in some distortion to calculated MREs in some cases and possible small downward skew in MRE in others. We are satisfied that the resulting changes are not material.
- Whilst the R code and spreadsheet model both have audits and internal checks, there are gaps in the recorded audit trail. However, we are comfortable that SDG have carried out the requisite manual checks across the gaps.
- Overall, we have confidence in the Network Rail payment rates calculated during this work.

We have made some recommendations for SDG to follow up with if there is a later re-run of the NRPR work:

- SDG should provide a full summary audit trail for each operator, showing the revenue in play at each stage of the full process from Lennon source data to flow-level MRE calculation, clearly identifying where losses occur and adjustments are made.

RDG SCH8 AUDIT - NRPR
REPORT

- SDG should consider correcting for any revenue lost through failure to map service groups, similarly to the way location mapping failures are corrected for.

And we have some recommendations for the future stages of the recalibration that will simplify the future audit work and improve confidence in the results for these stages:

- Logging from R code should be made more comprehensive and should include row counts to assist in detecting data duplication issues during merge operations.
- The R logs should be consolidated into a high-level audit trail spreadsheet to give a clearer view of data integrity and any losses across the entire data pipeline.
- Data consolidation / process control spreadsheets should log incoming and outgoing data totals and row counts.
- Data processing code and spreadsheet models should be tested using a set of test cases and independently-calculated expected results, similarly to the test harness created by VRC for the NRPR model.

1 Introduction

This document is Vivacity Rail Consulting's audit report on the calculation of Network Rail Payment Rates carried out by SDG as part of the PR18 Schedule 8 recalibration

1.1 BACKGROUND

RDG is managing the recalibration for Control Period 6 (CP6) of the parameters used in the performance regime formalised in Schedule 8 of the Track Access Agreement between Network Rail (NR) and each Train Operating Company (TOC). These parameters are:

- Network Rail Payment Rate (NRPR): the cost per day for each service group of the change in NR-caused weighted average lateness of 1 minute. This represents revenue presumed lost to the operator because of delays and cancellations to its trains caused by NR and other operators.
- Network Rail and Operator benchmarks (performance points): the expected levels of service group average lateness caused by Network Rail and the Operator. The Network Rail levels are set for each year of CP6 to be consistent with the regulatory targets NR will be set.
- TOC Payment Rate (TPR): the cost per day for each service group of the change in TOC-caused weighted average lateness of 1 minute. This recoups for Network Rail the liability it faces by delays caused by this service group to other operators.
- Sustained Poor Performance Threshold (SPP): the level of lateness caused by NR at which TOCs become eligible to recover additional amounts, under the assumption that the normal Schedule 8 entitlement is inadequate.

The bulk of the work in calculating these parameters is being carried out by Steer Davies Gleave (SDG), with the exception of a small number of operators where a significant service change is taking place, where the work is being done as bespoke recalibrations by other consultants.

1.2 THE AUDIT - TERMS OF REFERENCE

RDG require that the recalibration work be audited by a third party. The purposes of the audit are

- To verify that the correct input data have been used and flow correctly through the calculation models

RDG SCH8 AUDIT - NRPR
REPORT

- To check that any data manipulation is appropriate
- To check that the models used conform to the methodology agreed between SDG and the Schedule 8 Working Group
- To check that formulae in the models are correct
- To check that assumptions used in the models are valid and appropriate.

Vivacity Rail Consulting (VRC) have been engaged by RDG to carry out this audit. The VRC team have been involved in several previous Schedule 8 recalibrations, including the PR13 national recalibration for Control Period 5.

1.3 THIS REPORT

This report covers VRC's audit of the calculation of Network Rail Payment Rates. It is structured as follows:

- Section 2 describes our approach to the audit: the methodology and scope
- Section 3 describes our findings and conclusions
- Appendices show details of the work done.

2 Audit Approach

2.1 VRC'S OVERALL AUDIT APPROACH

VRC's approach to the audit task is three-stage, aiming at all times to pre-empt possible audit issues as early in the recalibration process as possible. The three stages are:

1. **Early engagement:** meeting with the modelling team to set out audit expectations and make suggestions about the model development process and data pipeline to a) minimise the likelihood of errors creeping in; b) simplify the later audit by providing a clear audit trail and simple model structure.
2. **Continuous monitoring:** regular contact with the modelling team to discuss emerging issues and apply any audit-related course corrections while they can still have an impact; early sight of models as they hit internal review points prior to formal delivery to maximise the time available to resolve any issues found.
3. **Formal audit:** review following draft delivery of the inputs, methodology, data pipeline and spreadsheet models.

2.2 NRPR MODELS AND PROCESS - APPROACH TO THE AUDIT

2.2.1 INTRODUCTION

By the time VRC had been appointed to carry out this audit, the NRPR model was nearly complete and was available in early draft form. It was thus too late to provide any effective input in terms of stages 1 and 2 above. The audit of the NRPR model has thus had to be retrospective and has focussed on the following areas:

1. Verification that the model methodology conformed with the referenced elements of the Passenger Demand Forecasting Handbook (PDFH)
2. Verification that the values and parameters in the relevant PDFH tables had been correctly transcribed into the model
3. A thorough check of the model calculations and logic, carried out by preparing a test harness and a set of test cases which exercised all the possible combinations of flow geography, distance, Generalised Journey Time (GJT), journey purpose, ticket type, peak period and airport involvement, as well as verified that the model formulae and ranges correctly covered all possible input data items.

4. A review of the upstream data processing pipeline which prepares Lennon revenue, assigns it to ticket type (Full/Reduced/Seasons/Other) and merges it with MOIRA-sourced geographical data to provide GJTs and distances.
5. A review of the code which populates and calculates the models for all the operators.

For all the audit steps listed above, more detail is provided in the paragraphs below, detailed results of our analysis are shown in the Appendices and the findings (including issues raised) are discussed in Section 0 below.

The audit team have reviewed the models for two operators (Chiltern and Northern) to verify that the processes which populate them and the assumptions made for them are correct. We have not reviewed models for the other operators.

2.2.2 OVERVIEW OF THE SDG PROCESS AND MODEL SUITE

The SDG process has essentially two stages. The first stage processes revenue data from Lennon into flows with known Generalised Journey Time (GJT), distance, geography and other characteristics used to calculate their Marginal Revenue Effect (MRE) - i.e. their contribution to the overall Network Rail payment rate - plus the associated revenue; the second calculates the MRE for each flow and then aggregates these to arrive at an overall Network Rail Payment Rate for each Service Group.

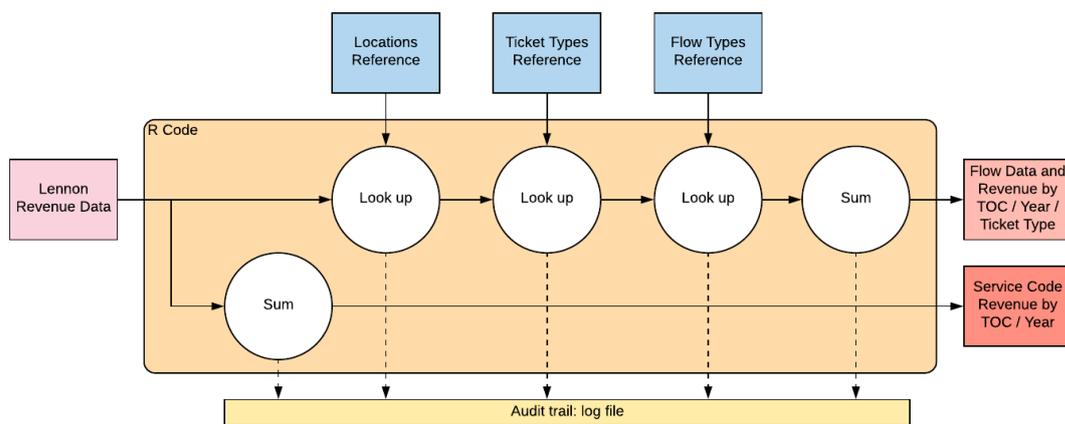


FIGURE 1 LENNON FLOW DATA PROCESSING

Figure 1 shows the structure of the first processing stage. Extracts of revenue data from Lennon for each operator and calibration year are

RDG SCH8 AUDIT - NRPR
REPORT

processed using a set of code modules written in the "R" statistical manipulation language to generate two outputs:

- A set of text files containing flow data for each year. There are actually two types of data in these files: flow-by-flow data for those flows responsible for most revenue; and data for the remaining flows, grouped up into GJT and distance bands.
- A set of text files containing totals of revenue by service code, extracted prior to the processing of flows. These totals are used as the service group level revenue and are used in the NRPR processing to scale the flow-level data back to the service group total to account for any discrepancies that may be introduced during the flow data processing.

The raw Lennon revenue data is augmented by a series of lookups into reference data:

- Location data is used to map the Lennon NLC codes for flow origin and destination to CRS codes used by MOIRA, and to identify the UK region that each location is in so that the geography of the flow can be identified.
- A ticket type lookup is used to map the Lennon product groups to the revenue categories Full / Reduced / Season used in the PDFH analysis.
- A flow type reference is used to characterise each flow according to the geographic region (London / South East / Rest of Country) of its origin and destination.

The R code creates a text log file as it runs, listing out the revenue, journeys and passenger miles handled at each stage of its processing.

The code creates a folder structure for its output, with files for each operator and year going into a different directory.

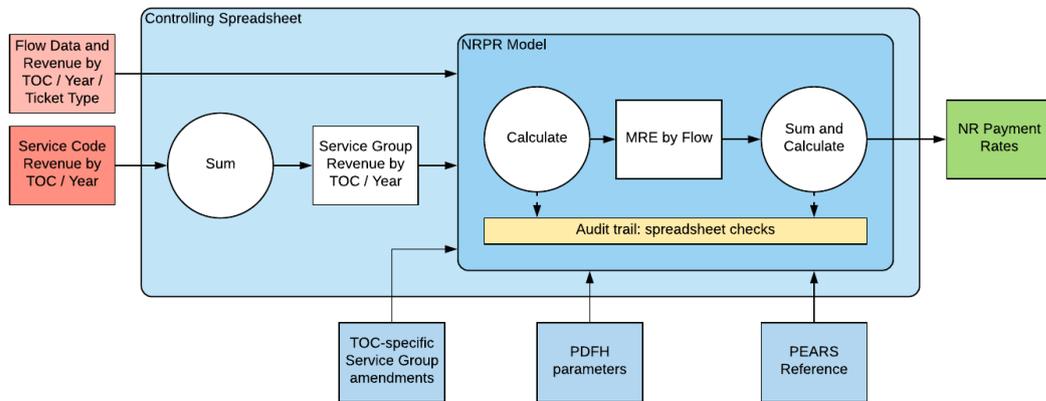


FIGURE 2 NRPR MODEL SPREADSHEETS

Figure 2 shows the structure of the second processing stage which calculates the NRPRs. This process is spreadsheet-based. A controlling spreadsheet iterates through a list of operators and carries out the following steps for each one:

- Creates an NRPR model spreadsheet from a template
- Populates the template with reference data common to all operators
- Loads in the PEARS data relevant to the operator
- Loads in the TOC-specific instructions for the operator: which years are being used for the calibration; any adjustments to the Lennon revenue agreed with the operator to cater for other revenue sources or expected changes due to, for example, timetable or fleet changes.
- Loads in the flow data for the operator, combining files where necessary
- Calculates totals of revenue for each PEARS service group and loads them to the model
- Runs the model macro which calculates and stores the MRE for each flow
- Makes the model spreadsheet calculate the NRPRs
- Saves the model spreadsheet for the operator.

The NRPR model for each operator has an internal audit trail in the form of a set of checks of data integrity and consistency.

RDG SCH8 AUDIT - NRPR
REPORT

3 Audit Review and Findings

3.1 INTRODUCTION

This section describes the findings of the audit steps described in Section 2. It concludes with a summary of our conclusions from the audit about the level of confidence stakeholders can have in the calculation of Network Rail Payment Rates.

During the course of the audit we have reviewed documents, code, data files and spreadsheets provided by SDG and RDG. [APPENDIX A](#) contains a complete list of these documents.

3.2 AUDIT OF THE DATA PREPARATION PIPELINE

3.2.1 WORK DONE

In our audit of the Data Preparation Pipeline, VRC have inspected:

- the R code which carries out the data preparation
- the input data files used as reference data, except for the locations reference data
- a sample log file of a run of the R code
- SDG's own internal review of the R code.

We have not reviewed or verified the correctness of the Lennon data extracts.

We have built a spreadsheet that parses the output log file to show the revenue figures, journeys and miles at each processing stage in a tabular form to allow us to look for any anomalies that might appear in the process.

At the time of writing we have not inspected the Locations lookup spreadsheet. This is the subject of a Clarification Question (CQ): Item https://github.com/VivacityRail/RDG_PR18_Ph3/issues/11.

3.2.2 OBSERVATIONS

The R code is of good quality and gives very little cause for concern. We have recorded some concerns about aspects of the reference data that may have a measurable distorting effect on flow-level MREs, and about the use

RDG SCH8 AUDIT - NRPR
REPORT

of the log file as an audit trail. We have also logged a single Severity 4 issue.

The issue concerns a formula in the part of the R code that looks up the GJT band for residual flows. This will fail to find a band where GJT is 1440 minutes or higher. (Issue https://github.com/VivacityRail/RDG_PR18_Ph3/issues/12). The impact of this on calculated MREs will be minimal.

Analysis of the process log file (see [APPENDIX C](#)) shows that the Locations lookup stage of the process is the one which introduces data variation. It is reasonable to expect that there would be a small amount of data loss here, as some Lennon locations fail to map to MOIRA station codes. This would be corrected later in the process as the flow-level revenue is scaled to match the service group total revenue. This is the case for most operators. However, we found some anomalies for which clarification questions were raised:

- For some service codes and for some operators, we see that the revenue increases rather than decreases. This suggests that there is some duplication of data occurring at this stage: this would happen if there were multiple matches in the reference data for a given Lennon flow. We have asked to see the Locations file so we can make our own assessment (CQ https://github.com/VivacityRail/RDG_PR18_Ph3/issues/11).
- For Merseyrail, there is a large drop in revenue (about 27% over the two years of the calibration) compared to the 1% or 2% typical for other operators (CQ https://github.com/VivacityRail/RDG_PR18_Ph3/issues/13, now [resolved](#)).

We believe the log file should include row counts as well as revenue and journeys totals. This would immediately reveal the presence of suspected multiple matches.

We should make clear that the data issues detected here do not invalidate the NRPR calculations. Their impact will be secondary and driven by whether there is a significant bias in the flow distance /GJT and ticket type in the affected flows compared to the operator's other flows.

We have not seen any evidence of SDG using the log file from this process as part of a published audit trail verifying that the revenue used to calculate the NRPRs matches that fed in to the start of this process. We have asked a CQ on this matter (https://github.com/VivacityRail/RDG_PR18_Ph3/issues/10). However, neither

RDG SCH8 AUDIT - NRPR
REPORT

do we have any reason to suspect they do not match. We believe such an audit trail to be desirable.

3.3 AUDIT OF THE CONTROLLING PROCESS

3.3.1 WORK DONE

VRC have inspected the structure, formulae and macro code of the spreadsheet model which controls the process of iterating through the operators and populating and running the model which calculates the NRPRs. The model is "SERVICE GROUP UPLIFTS AND PEARS V0.10.XLSB".

VRC have also reviewed some of the upstream models which process reference data from PDFH, PEARS and MOIRA to create the lookup data used by this model.

3.3.2 OBSERVATIONS

The controlling process model has been constructed broadly according to best-practice principles, though some output sheets contain calculations. We do not see this as a major weakness.

The macros are commented and clear to follow. There is nothing to suggest that this model will behave other than as intended.

We have seen some weaknesses:

- There is no Checks sheet in this model so no internal checking of the integrity of the data.
- There is a possibility of data loss in the mapping of Lennon service codes to PEARS Capri codes. The model makes an assumption that there is a direct mapping between the first 3 characters of the Lennon service code and the PEARS Capri code. In some cases no such mapping occurs. These are flagged in the spreadsheet as "SG Not Found". For the case of Scotrail, whose data was in the model we reviewed, this represented a 3.4% loss of revenue. This is at the higher end of what we would consider acceptable. A CQ was raised to see if this has been addressed in a more recent run or compensated for elsewhere ([https://github.com/VivacityRail/RDG PR18 Ph3/issues/15](https://github.com/VivacityRail/RDG_PR18_Ph3/issues/15), now [resolved](#)). The NRPR calculation spreadsheet model contains a check for this condition which will fail if the level of "SG Not Found" revenue exceeds 2.5%.
- There is no formal audit trail connecting the log file outputs from the R processing to the inputs and outputs of this model. This means that it is impossible to verify from observation that the correct

**RDG SCH8 AUDIT - NRPR
REPORT**

version of the Lennon data has been loaded and that the totals from the R code agree with those in this model.

3.4 AUDIT OF THE INPUT DATA TO THE NRPR CALCULATION MODEL

3.4.1 WORK DONE

VRC reviewed the inputs that Network Rail Payment Rate calculation model uses as reference data in the calculation of NRPRs. The inputs are:

- Semi-elasticities for London flows, from Oxera and the SDG peer review of this work
- GJT elasticities by geography and distance, from PDFH v6.0
- Delay multipliers by geography, distance, journey purpose, from PDFH 5.1
- Journey Purpose breakdown by geography, distance and ticket type, from PDFH 6.0, calculated in subsidiary spreadsheet model "PEAK PROFILES AND JP_TT MAPPING V0.08.XLSX".
- Peak / Off-Peak split by distance, journey purpose and peak station "blueness", from MOIRA, calculated in subsidiary spreadsheet model "PEAK PROFILES AND JP_TT MAPPING V0.08.XLSX".
- Peak day-of-week splits by peak location, distance band and ticket type.
- Airport traffic proportions for airport flows, from the CH2M PR13 recalibration.
- Operator-specific overrides at the Service Group level, including which years' data should be used in the calibration, and any adjustments to the Service Group level revenues to cater for non-Lennon sources or reasons that future revenue might differ from past such as significant changes to the timetable or fleet.

The Oxera and PDFH data items are nearly all direct entries from the original sources. Here, we have checked that the values have been correctly transcribed into the model. Where this is not so, some calculation has been done, as follows:

- The Journey Purpose breakdown from PDFH 6.0 is rescaled from the original input to suit the method of calculation. We have checked the calculation.

RDG SCH8 AUDIT - NRPR
REPORT

- The Peak / Off-Peak split is calculated by aggregating data from MOIRA 15-minute profiles. We have checked the calculation but have not checked the original source of the MOIRA profile data.

We have not been able to check the calculation of the Peak day-of-week splits as the source of the table in the NRPR spreadsheet is not specified and the data lineage provided so far is not clear. We have raised a CQ (https://github.com/VivacityRail/RDG_PR18_Ph3/issues/14) for further information on this point.

3.4.2 OBSERVATIONS

All the Oxera and PDFH parameters appear to have been correctly passed into the NRPR model.

The spreadsheet model "PEAK PROFILES AND JP_TT MAPPING V0.08.XLSX" has been constructed using an SDG template. However, the standard of construction is not as high as for the other models examined, in the following regards:

- It contains several input sheets which are not used
- It does not have a separate calculation and outputs section, so it is not clear which sheets represent actual outputs
- It has an untitled and mislabelled calculation block on one of the input tabs.
- The version history block is not up-to-date.

These notes aside, the calculations that lead to the presumed output sheets appear to have been done correctly.

The Journey Purpose re-scaling calculation appears to have been done correctly.

The calculation of Peak / Off-Peak split is based on MOIRA demand profiles which are defined in terms of 15-minute periods during the day for each combination of From / To London, From / To "Blue" (=important) station / From/To Other station, direction of travel (Outward/Return), peak station and journey purpose. The calculation appears to have been done correctly.

The spreadsheet model has a tab "IC; Day of Week" which seems to be intended to calculate the proportion of travel (i.e. revenue) which occurs on days of the week which have peaks. However, the sheet is not labelled, no data source is identified and the output block which appears on the tab does not correspond to the day-of-week split table input to the NRPR calculation spreadsheet. As noted above, we have raised a CQ to get more information about this split.

**RDG SCH8 AUDIT - NRPR
REPORT**

It should be noted that any discrepancy in the calculation of Peak / Off-Peak or Day-of-week splits will only move MRE between the Peak and Off-Peak divisions of the same service group, raising one and lowering the other. The operator-level and industry impact would therefore be minimal.

Overall we can be comfortable that the NRPR calculation model is using the correct input lookup data.

3.5 AUDIT OF THE MODEL METHODOLOGY

3.5.1 WORK DONE

VRC have carried out a careful review of the method used in the calculation of flow-level Marginal Revenue Effects and the aggregation of these to the service group level to calculate Network Rail Payment Rates. The purpose of the review was to verify that the model correctly implemented the methodology agreed by SDG and the Schedule 8 Working Group. It was not to comment on the validity of this methodology.

The work took particular note of the following methodological points:

- Use of the modified derivative-based formula which looks at the slope of the MRE function at the actual GJT for each flow, rather than the impact of adding 1 minute to the GJT.
- Application of the correct formula for MRE depending on whether the flow was a London one or not.
- Handling of the Peak day-of-week and Peak / Off-Peak proportions
- Handling of Airport flows, especially given the fact that flows could start or end at an Airport
- Coverage of residual flows - those where individual origins / destinations are grouped by distance / GJT / geography.
- Correction of revenue lost through flow mapping.

3.5.2 OBSERVATIONS

Once a number of clarification questions had been addressed, we observed no difference between the working of the model and the agreed methodology. We can be comfortable that the NRPR model is "doing the right calculation" - i.e. it is attempting to use the correct method.

APPENDIX B lists the methodology points mentioned in the SDG methodology document and notes the extent to which we have verified them.

3.6 DETAILED AUDIT OF MODEL FUNCTIONALITY VERSUS EXPECTED RESULTS

3.6.1 WORK DONE

Since VRC were not able to get involved early in the NRPR model process, we determined that we should give the NRPR model a particularly thorough test once it was provided to us. We did this by setting up a set of dummy flows which between them covered all the permutations in the various PDFH data tables, as well as verifying that the spreadsheet calculations correctly covered all the input data ranges.

For each dummy flow, we independently calculated the expected MRE using our own interpretation of the intended methodology, then compared our results with those calculated by the NRPR model.

We also carried out our own calculation of Network Rail payment rates by aggregating the flow-level MREs and compared results with those calculated by the NRPR model.

Finally, we prepared a set of "pathological" flows, with various data quality failings, to see whether the NRPR model would detect them and throw an error, so alerting users of the presence of the problem.

To ensure coverage of all the possible data cases, we set up a series of decision tables for each of the questions needing to be asked of a flow to specify how its MRE should be calculated. The decision tables were:

- Table 1: which type of calculation to use - London or Rest of Country - based on flow geography. 6 cases.
- Table 2: London calculation: choice of semi-elasticity, based on geography and ticket type. 6 cases.
- Table 3: Rest of Country calculation: identification of PDFH distance group, Journey Purpose distance group, Sector and GJT elasticity based on flow geography and distance. 17 cases.
- Table 4: Rest of Country calculation: identification of Delay Multiplier based on geography, distance, ticket type, journey purpose and airportness. 131 cases.
- Table 5: Rest of Country calculation: identification of Peak day percent and Peak / Off-Peak split based on ticket type, GJT band, peak location and journey purpose. 244 cases.
- Table 6: Rest of Country calculation: identification of Journey Purpose proportion based on Geography, Distance, and Ticket Type. 162 cases.

**RDG SCH8 AUDIT - NRPR
REPORT**

We created a test case for each of the decision table cases listed here. This was done using spreadsheet formulae to look up into the decision tables and create flow data with the appropriate characteristics.

We created a further set of test cases which would check that the spreadsheet formulae were referring to the correct data ranges. These cases had data in all the "corners" of the data ranges: first and last row, first and last column. They would pick up any misalignment of data ranges in the spreadsheet model.

Lastly, we created a set of "pathological" flows: ones with deliberate errors in them that would check that the model could not pass invalid data through its calculations without throwing an error or triggering one of its internal checks.

To support the flow data, we created a bespoke set of lookup data which had different values in every cell, or was set up to guarantee that only a single journey purpose or peak status would be involved. This would enable us to check that the correct values were being selected from the lookup tables for each flow type. The actual reference data has many repeats of the same value in different cells, so might give a matching result even though it was performing an incorrect calculation.

With these data we calculated flow MREs and Service Group NRPRs with our own test harness - a mock-up of the model using our own interpretation of the intended methodology. We also ran the actual NRPR model and compared the results with those from our mock-up.

APPENDIX D shows extracts of the decision tables and some of the mock-up's flow calculations.

APPENDIX E shows the correspondence at the flow level between the mock-up flow level calculations and the NRPR model for a recent test run.

3.6.2 OBSERVATIONS

We found exact agreement between our test harness mock-up in 226 of the 234 test cases tried. Of the 8 where there was a discrepancy, this was caused as follows:

- 4 cases revealed an error in the way the test harness was calculating MREs for short-distance London-based flows involving airports
- 3 cases were invalid test cases in which the data had been set up wrongly
- 1 case was caused by differences in the reference data between test harness and model

RDG SCH8 AUDIT - NRPR
REPORT

None of these differences can be attributed to a fault in the NRPR model - they require more work on our part to correct the test harness logic and the test data.

Where we created pathological flows, these were detected by the spreadsheet model and resulted in errors (#N/A, #VALUE or #DIV/0) which would indicate their presence and lead to correction of the input data.

3.7 SUMMARY OF ISSUES RAISED

During the course of the audit, we have raised issues with SDG using our on-line issue tracking tool based on Github. This is hosted at https://github.com/VivacityRail/RDG_PR18_Ph3/issues; access has been granted so far to David Ford at SDG and Caitlin Scarlett at RDG. The issues are summarised in **TABLE 1** below.

TABLE 1 ISSUES RAISED ON NRPR AUDIT

Issue	Description	Severity	Status	Note
2 ¹	Flow level revenue higher than SG level	3	Closed	Linked with Item 11
3	GJT set to 1400 mins leads to low MRE	3	Closed	No significant impact on MRE. Initial occurrences corrected in data.
4	Bidirectional Flows?	CQ	Closed	
5	Peak Profiles and JP_TT mapping	CQ	Closed	
6	Interface between R code and spreadsheet	CQ	Closed	
7	Handling of other revenue	CQ	Closed	
8	20000-row cutoff for flow-level analysis	CQ	Closed	

¹ Issue 1 was a test issue used to demonstrate the online issue tracker

RDG SCH8 AUDIT - NRPR
REPORT

Issue	Description	Severity	Status	Note
9	Airport revenue percentage: commuting	CQ	Closed	
10	R pipeline - log file	CQ	Closed	Request for update on log file
11	R pipeline - Module 4 data and diagnostics	CQ	Closed	Data request to enable VRC to check data mismatches
12	R code - residual GJT lookup	4	Closed	Very minor issue with minimal impact
13	R pipeline - Merseyrail loss in location match	CQ	Closed	
14	Peak proportions - day of week	CQ	Closed	Extra information requested to allow completion of audit. No evidence of model or data error at this stage.
15	Service groups and Uplifts - SG not found	CQ	Closed	Corrected in data - issues mostly associated with service groups split between operators.

No issue higher than a severity level 3 was found.

3.8 CONCLUSIONS AND RECOMMENDATIONS

3.8.1 CONCLUSIONS - NRPR CALCULATION

From the audit work, we can draw the following conclusions about the process, the code and models and the fitness for purpose of the resulting Network Rail Payment Rates:

- The R code and spreadsheet models have generally been built to a high standard
- The methodology used is sound and as agreed with the Schedule 8 Working Group

RDG SCH8 AUDIT - NRPR
REPORT

- No major issues have been found with the input data, model and calculation
- There were minor areas of concern to do with the location mapping and service group lookup which resulted in some distortion to calculated MREs in some cases and possible small downward skew in MRE in others. We are satisfied that the resulting changes are not material.
- Whilst the R code and spreadsheet model both have audits and internal checks, there are gaps in the recorded audit trail. However, we are comfortable that SDG have carried out the requisite manual checks across the gaps.
- Overall, we have confidence in the Network Rail payment rates calculated during this work.

3.8.2 RECOMMENDATIONS

We have the following recommendations.

For the Network Rail Payment Rates:

- SDG should provide a full summary audit trail for each operator, showing the revenue in play at each stage of the full process from Lennon source data to flow-level MRE calculation, clearly identifying where losses occur and adjustments are made.
- SDG should consider correcting for any revenue lost through failure to map service groups, similarly to the way location mapping failures are corrected for.

For future stages of the Schedule 8 recalibration work:

- Logging from R code should be made more comprehensive and should include row counts to assist in detecting data duplication issues during merge operations.
- The R logs should be consolidated into a high-level audit trail spreadsheet to give a clearer view of data integrity and any losses across the entire data pipeline.
- Data consolidation / process control spreadsheets should log incoming and outgoing data totals and row counts.
- Data processing code and spreadsheet models should be tested using a set of test cases and independently-calculated expected results, similarly to the test harness created by VRC for the NRPR model.

Appendix A Items Reviewed

The tables in this Appendix list the items reviewed as part of the audit.

A.1 SOURCE DOCUMENTATION AND METHODOLOGY

The documents in **TABLE 2** were reviewed and compared against the methodology actually used in the NRPR model spreadsheet. The same methodological principles were used in our test harness which independently calculated flow MREs for comparison with the NRPR model.

Where these documents provided reference data, we checked that the data had been correctly transcribed into the model.

TABLE 2 REVIEW ITEMS - BACKGROUND AND METHODOLOGY DOCUMENTATION

Item	Document reference	Relevance
SDG NRPR Methodology document	Recalibration of Schedule 8 for CP6 - Methodology (INITIAL Draft) v0.025.docx	Definition of the methodology to be followed by SDG. Original version released to auditors in March 2018
	Recalibration of Schedule 8 for CP6 - Methodology v0.30 (Final Draft of NRPR approach).docx	Updated version released to auditors on 29/5/2018
PDFH v5.1	B Combined.pdf	Source for Delay Multipliers for different geography / journey purpose combinations
PDFH v6.0	PDFH6B1IntroductionandMarketSegmentation.pdf	Introduction and background to PDFH 6.0 approach
	PDFH6B4GJT.pdf	Source for Generalised Journey Time elasticities
Oxera	170830 The impact of unplanned disruption on	Study proposing the

RDG SCH8 AUDIT - NRPR
REPORT

Item	Document reference	Relevance
analysis	train operator revenue FINAL.pdf	use of semi-elasticities for heavily commuter-focused flows
SDG review of Oxera analysis	SDG Unplanned Disruption Peer Review Report v1.0.pdf	Peer review of the Oxera study, source of the semi-elasticities to be used for London-London and London-South East flows.
Arup MRE formulation	MRE formulation paper.pdf	Proposal of the derivative formulation for the change in MRE per unit change in GJT
Halcrow CP5 report	halcrow-sch8-recalibration-2013-10-07.pdf	CP5 methodology; airport travel percentages for airport flows

A.2 DATA PIPELINE ELEMENTS

The items in **TABLE 3** were reviewed to verify that the data processing pipeline which prepares Lennon flow data is sound.

TABLE 3 REVIEW ITEMS - DATA PROCESSING PIPELINE

Item	Document reference	Relevance
R code	NRPR_master_03042018 .R	Code in the R language used to process the Lennon revenue data by flow ready for use in the NRPR spreadsheet
SDG internal review of R code	Review of R Code (NR Payment Rates).docx	Enables VRC to understand quickly the layout of the R code and assess the quality of SDG's internal quality control processes.
Lennon flows extracts	London Midland_yyyy.csv and .txt; Transpennine_yyyy.csv and .txt; VT East Coast-yyyy.csv	Sample flow-level revenue files used to investigate whether issues identified in the NRPR model for one operator also affected other operators

RDG SCH8 AUDIT - NRPR
REPORT

Item	Document reference	Relevance
	and .txt	
MOIRA GJT reference data	MOIRA_WITH_BR_v2.csv	Data file which gives distance for each origin-destination pair of stations, plus GJTs for Full, Reduced and Standard ticket types. Used to provide these data items to Lennon flows. This file has over 6.5m rows
Flow types reference	flows_lookup.csv	Maps origin and destination location areas (S_LN, S_SE, S_OSE) to PDFH geographical flow types
Ticket types reference	Ticket Lookup.xlsx	Maps Lennon product group to the MOIRA / PDFH ticket types Full / Reduced / Seasons / Other
Locations reference	Improved Station Lookup vxx.xlsx	Locations reference - maps Lennon flow origin and destination to BR stations suitable for lookup in MOIRA data. <Note - file requested but not yet reviewed. Subject of CQ https://github.com/VivacityRail/RDG_PR18_Ph3/issues/11 >
Log File	LogFile_2018-03-29_10_19_31.txt	Log file produced by run of the R code on a full set of data. Analysed to assess where mapping issues cause data inconsistencies.

A.3 TOC-SPECIFIC DATA PREPARATION AND MODEL CONTROL

We reviewed the items in **TABLE 4**. These are associated with the spreadsheet model which controls the process of loading the correct data for each operator's NRPR payment rate calculation, running the calculation and saving the model.

TABLE 4 REVIEW ITEMS - CONTROLLING PROCESS AND TOC-SPECIFIC DATA

Item	Document reference	Relevance
Controlling model	Service Group Uplifts and PEARS v0.10.xlsb	Spreadsheet which manages the process of creating and populating the NRPR model for

RDG SCH8 AUDIT - NRPR
REPORT

Item	Document reference	Relevance
		each TOC, including loading its TOC-specific adjustments
Peak Profiles Calculation	Peak Profiles and JP_TT Mapping v0.08.xlsx, Calibrated Day of Week Splits.csv	Spreadsheet and data file which estimate the proportion of revenue associated with Peak and Off-Peak sub-service groups based on MOIRA demand and day-of-week profiles and PEARS peak times.

A.4 NRPR MODEL

The items listed in Table 5 are versions of the NRPR model which we reviewed at different stages in its lifecycle.

TABLE 5 REVIEW ITEMS - NRPR MODEL

Item	Document reference	Relevance
Early draft versions of NRPR model	Network Rail Payment Rates Model v0.046 (Initial Draft) - Review.xlsm, Network Rail Payment Rates Model v0.047 (Initial Draft) - Review.xlsm	Spreadsheet model which calculates flow-level MRE and aggregates to service group Network Rail Payment Rates
Released draft version of the NRPR model (Chiltern)	Chiltern - Network Rail Payment Rates Model v0.20 (Second Draft).xlsm	Model updated with revised handling of airport flows and improvements to upstream data mapping
Updated version of the model (Northern)	Northern Rail - Network Rail Payment Rates Model v0.20 (Second Draft).xlsm	Model updated with further revision to airport flows and inclusion of additional data integrity checks; check of non-London operator.
TOC-specific notes - Northern	Northern - NRPR TOC-Specific Notes and Version Control v0.20 (Second Draft).docx	Description of the TOC-specific data overrides applied.
SDG internal review document on NRPR model	Steer Davies Gleave Data Flows Report v0.46.xlsx	Enables us to assess the thoroughness and quality of the SDG review process and identify any areas that our audit should focus on

RDG SCH8 AUDIT - NRPR
REPORT

Item	Document reference	Relevance
SDG internal review of model structure and formulae	OAK Summary v0.46.xlsx	Enables us to see quickly the structure of the model and the types of formulae employed in in it.

Appendix B Methodology Checks

In this Appendix we list the points of methodology that VRC have checked to have been correctly handled in the NRPR calculation model suite. In **TABLE 6**, the Reference column refers to the SDG document "RECALIBRATION OF SCHEDULE 8 FOR CP6 - METHODOLOGY V0.30 (FINAL DRAFT OF NRPR APPROACH).DOCX" and the Checked In column refers to the item(s) in the modelling suite where we verified the application of the methodology point.

In the Note column, absence of a note means we have confirmed by inspection that the methodology point has been applied in the calculation. A note may refer either to a more in-depth test that VRC have done, an indication of any aspects that we have not checked, or any discrepancies we have observed.

TABLE 6 METHODOLOGY CHECKS

Reference	Methodology point	Checked in	Note
2.4	NRPR formula	NRPR model, source documentation	Verified by independent test
2.5	MRE formula - SE and London	NRPR model, source documentation	Verified by independent test
Table 2.1	Semi-elasticities	Control spreadsheet, NRPR model, source documentation	Verified by independent test
2.6	MRE formula - GB not London / SE	Control spreadsheet, NRPR model, source documentation	Verified by independent test
Table 2.2	Delay multipliers	Control spreadsheet, NRPR model, source	Verified by independent test

RDG SCH8 AUDIT - NRPR
REPORT

Reference	Methodology point	Checked in	Note
		documentation	
Table 2.3	GJT elasticities	Control spreadsheet, NRPR model, source documentation	Verified by independent test
2.7	Demand data processing	R code	
2.8 Table 2.4	1 Lennon revenue files	Lennon data files	Sample files for various TOCs spot checked.
	2 Ticket category lookup	R code module 3, Ticket Lookup.xlsx	
	3 NLC and TLC location lookup	R code module 4	Locations lookup file not yet reviewed - subject of CQ https://github.com/VivacityRail/RDG PR18 Ph3/issues/11
	3 GJT / Distance lookup	R code module 6, MOIRA_WITH_BR_v2.csv	
	4 Area lookup		Part of locations lookup - see CQ above
	4 Sector lookup	R code module 5, flows_lookup.csv	
	5 Journey Purpose	Control spreadsheet, NRPR model, Peak Profiles and JP_TT Mapping v0.08.xlsx	Methodology document suggests done in demand data pipeline. In fact done in NRPR spreadsheet model. Verified by independent test

RDG SCH8 AUDIT - NRPR
 REPORT

Reference	Methodology point	Checked in	Note
	6 Peak Proportion	Control spreadsheet, NRPR model, Peak Profiles and JP_TT Mapping v0.08.xlsx	<p>Methodology document suggests done in demand data pipeline. In fact done in NRPR spreadsheet model.</p> <p>Model operation checked and verified by independent test. Source data for Peak/Off-Peak mapping checked; source data for Day of Week mapping not checked - CQ https://github.com/VivacityRail/RDG_PR18_Ph3/issues/14 raised for this.</p>
2.9	Lennon data extract	Various Lennon text extract files	
2.10	Aggregation to ticket category	R code module 3, Ticket Lookup.xlsx	<p>Steps are in different order in Methodology document than in R code.</p> <p>See entry under Table 2.4 above</p>
2.11	Resonate GJT lookup		See entry under Table 2.4 above
2.12	Location lookup		See entry under Table 2.4 above
2.13	Sector		See entry under Table 2.4 above
2.14	Area		See entry under Table 2.4 above
2.15	Sector to MRE Method map		See entry under Table 2.4 above

RDG SCH8 AUDIT - NRPR
REPORT

Reference	Methodology point	Checked in	Note
2.16	Distance bands	R code, module 7	Spot checks of distinct distance values vs bands to verify correct banding
2.17	Journey purpose allocation		See entry under Table 2.4 above
2.18	Peak / Off-Peak allocation		See entry under Table 2.4 above
2.19	Adjustments - timeframe (also Appendix 1)	Source data in tab I; Timeframe of control spreadsheet; macro to copy to TOC NRPR model; NRPR model	Handling of input data in NRPR model validated by independent test. Control macro checked by inspection of ranges and boundaries used.
	Adjustments - non-Lennon revenue	Source data in tab I; Timeframe of control spreadsheet	Note that the SG level uplifts can be applied for this purpose and some are noted as having been used this way. No verification that the uplifts are correct. Verified that the NRPR model applies the uplifts correctly by means of independent test.
	Adjustments - non-geographic revenue	Source data in tab I; Timeframe of control spreadsheet	We note that the service group uplifts mechanism can be used for this purpose and that it appears to have done so. We cannot verify the uplifts for any service

RDG SCH8 AUDIT - NRPR
 REPORT

Reference	Methodology point	Checked in	Note
			<p>group are correct.</p> <p>We have verified by means of independent test that uplifts defined here are correctly applied in the NRPR model.</p>
	Adjustments - peak/off-peak split	Tab I; Peak Split in the Peak Profiles Calculation spreadsheet	<p>We note an adjustment has been made for Cardiff. The adjusted figures have been passed through to the template NRPR model input sheets.</p> <p>We have verified by independent test that these profiles are correctly applied in the NRPR model calculation.</p>
	Adjustments - service code allocation	Tab I; Capri in the control spreadsheet which contains the Lennon service code to PEARS SG mapping	<p>The version of the control spreadsheet we have reviewed has a formula error which prevents the service codes being passed to the NRPR model correctly. CQ https://github.com/VivacityRail/RDG PR18 Ph3/issues/16 raised.</p>
	Adjustments - airport flows	Control spreadsheet I; Airport. NRPR model	<p>Airport flow % passed correctly to NRPR model and handled correctly in the calculation.</p> <p>Verified by independent test.</p>
	Adjustments - refunds and	Source data in tab I; Timeframe	We note that the service group uplifts mechanism

RDG SCH8 AUDIT - NRPR
REPORT

Reference	Methodology point	Checked in	Note
	non-issues	of control spreadsheet	can be used for this purpose.
	Adjustments - non-marginal revenue	Source data in tab I; Timeframe of control spreadsheet	We note that the service group uplifts mechanism can be used for this purpose.
2.20	Airport flow proportion	Tab C; Flow in NRPR model	Correct operation verified by independent test.
2.21	Airport flows - MRE calculation method	Tab C; Flow in NRPR model	Correct operation verified by independent test.
2.22	Residual flows - grouping	R code, module 9. Tab C; Flow in NRPR model	Handling of grouped flows verified by independent test.
2.23, Appendix 7	Residual flows - GJT band widths	R code, module 9	
2.24, Appendix 7	Residual flows - calculation method	R code, module 9	
Appendix 2	Price base factor	Tabs I; Price Base, C; SG MRE 2016 and C; SG MRE 2017 in NRPR model	Handling of entered price base correction factors verified by independent test.
A.2 Table A2.1	Factors	Tab I; Price Base, in NRPR model	Value for 2015/16 is 1.02143 in NRPR model, 1.02142 in Methodology document. Our calculation from https://www.ons.gov.uk/economy/inflationandpricein

RDG SCH8 AUDIT - NRPR
 REPORT

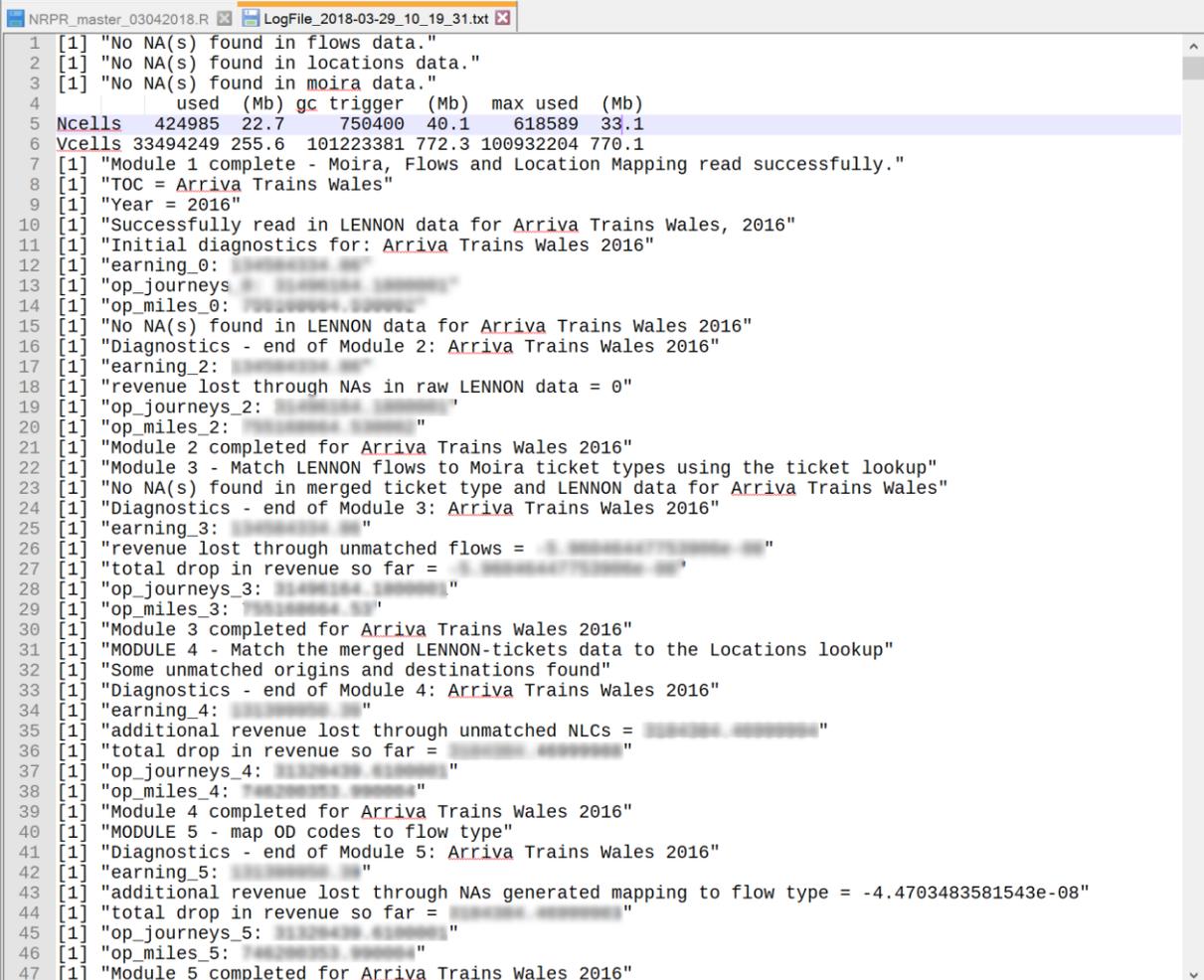
Reference	Methodology point	Checked in	Note
			dices/timeseries/chaw for Nov 16 / Nov 15 gives a value of $265.5/259.8 = 1.02194$. The difference is not significant
A.5	Use of RPI	Tab I; Price Base, in NRPR model	RPI has been used
Appendix 4	Busyness factors	Calculation of SG busyness in control spreadsheet, tabs I; Busyness, C; Busyness Annual and O; Busyness Annual. Application in NRPR model, tabs I; Busyness and C; Payment Rates.	Calculation from raw PEARS input and transfer to NRPR model appears correct. Application verified by independent test.
Appendix 5	Semi-elasticities	NRPR model, tab I; Parameters	Application in London and L-SE calculation verified by independent test
Table A5.2	Semi-elasticities - London and SE flows	NRPR model, tab I; Parameters	Values used agree with Table 8 on p9 of the SDG peer review of Oxera report.
Appendix 6 A.15 Fig A6.2	MRE equation - SE and London	Tab C; Flow in NRPR model	Verified by independent test
A.17 Fig A6.4	MRE equation - rest of GB	Tab C; Flow in NRPR model	Verified by independent test
Appendix 7 A.18 Table	Residual flows - GJT band	R code, module 9	

RDG SCH8 AUDIT - NRPR
REPORT

Reference	Methodology point	Checked in	Note
A7.1	widths		
Appendix 8 A.24 Table A8.1	Identification of "South East" stations	R code, module 4	Locations lookup file not yet checked - CQ https://github.com/VivacityRail/RDG_PR18_Ph3/issues/11
Appendix 9 A.25 Table A9.1	Peak locations, times and service groups	Peak Profiles calculation spreadsheet	
A.26	Cardiff peak days	Peak Profiles calculation spreadsheet	
A.28	Use of Calibrated Demand Profiles	Peak Profiles calculation spreadsheet	
	Use of Calibrated Day of Week Splits	Calibrated Day of Week Splits.csv; I; Day of Week tab in NRPR model	Exact derivation of the input not clear. CQ https://github.com/VivacityRail/RDG_PR18_Ph3/issues/14 refers
A.29	Adjustment for journey length	Peak Profiles calculation spreadsheet	Table in I; Name Type gives time offset in minutes for each band.

Appendix C Data pipeline log

The demand data processing pipeline logs its actions in a text file. **FIGURE 3** shows an example (with the TOC-specific figures obscured).



```

1 [1] "No NA(s) found in flows data."
2 [1] "No NA(s) found in locations data."
3 [1] "No NA(s) found in moira data."
4      used (Mb) gc trigger (Mb) max used (Mb)
5 Ncells  424985 22.7   750400 40.1   618589 33.1
6 Vcells 33494249 255.6 101223381 772.3 100932204 770.1
7 [1] "Module 1 complete - Moira, Flows and Location Mapping read successfully."
8 [1] "TOC = Arriva Trains Wales"
9 [1] "Year = 2016"
10 [1] "Successfully read in LENNON data for Arriva Trains Wales, 2016"
11 [1] "Initial diagnostics for: Arriva Trains Wales 2016"
12 [1] "earning_0: [REDACTED]"
13 [1] "op_journeys [REDACTED]"
14 [1] "op_miles_0: [REDACTED]"
15 [1] "No NA(s) found in LENNON data for Arriva Trains Wales 2016"
16 [1] "Diagnostics - end of Module 2: Arriva Trains Wales 2016"
17 [1] "earning_2: [REDACTED]"
18 [1] "revenue lost through NAs in raw LENNON data = 0"
19 [1] "op_journeys_2: [REDACTED]"
20 [1] "op_miles_2: [REDACTED]"
21 [1] "Module 2 completed for Arriva Trains Wales 2016"
22 [1] "Module 3 - Match LENNON flows to Moira ticket types using the ticket lookup"
23 [1] "No NA(s) found in merged ticket type and LENNON data for Arriva Trains Wales"
24 [1] "Diagnostics - end of Module 3: Arriva Trains Wales 2016"
25 [1] "earning_3: [REDACTED]"
26 [1] "revenue lost through unmatched flows = -5.0000000000000000e-08"
27 [1] "total drop in revenue so far = -5.0000000000000000e-08"
28 [1] "op_journeys_3: [REDACTED]"
29 [1] "op_miles_3: [REDACTED]"
30 [1] "Module 3 completed for Arriva Trains Wales 2016"
31 [1] "MODULE 4 - Match the merged LENNON-tickets data to the Locations lookup"
32 [1] "Some unmatched origins and destinations found"
33 [1] "Diagnostics - end of Module 4: Arriva Trains Wales 2016"
34 [1] "earning_4: [REDACTED]"
35 [1] "additional revenue lost through unmatched NLCs = [REDACTED]"
36 [1] "total drop in revenue so far = [REDACTED]"
37 [1] "op_journeys_4: [REDACTED]"
38 [1] "op_miles_4: [REDACTED]"
39 [1] "Module 4 completed for Arriva Trains Wales 2016"
40 [1] "MODULE 5 - map OD codes to flow type"
41 [1] "Diagnostics - end of Module 5: Arriva Trains Wales 2016"
42 [1] "earning_5: [REDACTED]"
43 [1] "additional revenue lost through NAs generated mapping to flow type = -4.4703483581543e-08"
44 [1] "total drop in revenue so far = [REDACTED]"
45 [1] "op_journeys_5: [REDACTED]"
46 [1] "op_miles_5: [REDACTED]"
47 [1] "Module 5 completed for Arriva Trains Wales 2016"

```

FIGURE 3 SAMPLE R CODE LOG FILE

VRC built a parser in MS-Excel for the log files to enable the revenue and journeys counts at each stage of the demand data processing pipeline to be summarised and the source of any mismatches identified.

RDG SCH8 AUDIT - NRPR
REPORT

2017	Stage 2 NA	Stage 3 TT	Stage 4 NLC	Stage 5 Flow	Stage 6 MOIRA	MOIRA Full	Stage 8	Residual / pre	Residual / post	Other	Overall mismatch in process	Mismatch %
Arriva Trains Wales												2.3%
C2C												1.0%
Caledonian Sleeper												-1.3%
Chiltern												1.6%
CrossCountry												2.2%
Crossrail												0.7%
East Midlands Trains												1.9%
Grand Central												1.6%
Great Western 1												1.5%
Great Western 2												0.6%
Great Western 3												3.8%
Greater Anglia												1.2%
Hull Trains												1.5%
Island Line												4.8%
London Midland												3.4%
London Overground												0.3%
Merseyrail												28.8%
Northern 1												1.4%
Northern 2												2.7%
Northern 3												2.0%
Northern 4												2.1%
Scotrail 1												2.2%
Scotrail 2												3.8%
South Western 1												0.3%
South Western 2												-1.4%
South Western 3												2.1%
Transpennine												2.1%
VT East Coast												3.0%
VT West Coast 1												1.8%
VT West Coast 2												1.1%

FIGURE 4 TABULATED REVENUE BY PROCESSING STEP

C.1 REVENUE TOTALS BY STAGE AND OVERALL MISMATCH

Figure 4 shows the of revenue totals processed at each stage for each operator, for 2017. The revenue figures have been obscured. 2016 is similarly tabulated. From this table, we see two immediate issues which were raised with SDG and investigated further:

- For two sets of data - Caledonian Sleeper and South Western 2 - the data pipeline managed to increase the revenue. This suggests some data duplication.
- For Merseyrail, the level of mismatch was very high.

The impact of any mismatch in this processing is secondary, as corrections are applied in the NRPR spreadsheet to re-scale each service group's allocated revenue back to the Lennon total revenue. MREs will be affected only to the extent that any mis-matched flows have significantly different characteristics overall from the matched flows.

C.2 MISMATCHES BY STAGE

2017	2 to 3	3 to 4	4 to 5	5 to 6	6 to M	M to 8
Arriva Trains Wales						
C2C						
Caledonian Sleeper						
Chiltern						
CrossCountry						
Crossrail						
East Midlands Trains						
Grand Central						
Great Western 1						
Great Western 2						
Great Western 3						
Greater Anglia						
Hull Trains						
Island Line						
London Midland						
London Overground						
Merseyrail						
Northern 1						
Northern 2						
Northern 3						
Northern 4						
Scotrail 1						
Scotrail 2						
South Western 1						
South Western 2						
South Western 3						
Transpennine						
VT East Coast						
VT West Coast 1						
VT West Coast 2						

FIGURE 5 LOSSES BY PROCESSING STAGE

We inspected the logs further to identify where mismatches were occurring in the data pipeline. From Figure 5 we can see that all significant change was taking place at Stage 4, where Lennon locations from the flow data are matched with BR stations in the MOIRA flow list.

Data changes at this stage can have 2 causes:

- Failure to match - where there is no O-D pair in MOIRA that matches the Lennon pair. This will cause a reduction in row count and revenue.

RDG SCH8 AUDIT - NRPR
REPORT

- Duplicate matches - where a Lennon O-D pair ends up matching more than one set of locations in the lookup data. This will tend to cause an increase in row count and revenue.

Following this investigation, the following issues were raised:

- https://github.com/VivacityRail/RDG_PR18_Ph3/issues/2 Severity 3: Flow level revenue higher than SG revenue
- https://github.com/VivacityRail/RDG_PR18_Ph3/issues/10 CQ: request for current version of log file
- https://github.com/VivacityRail/RDG_PR18_Ph3/issues/11 CQ: request for the locations lookup data file.
- https://github.com/VivacityRail/RDG_PR18_Ph3/issues/13 CQ, closed: Merseyrail data loss.

C.3 RESIDUAL FLOWS

The log file records that the processing of residual flows leads to a small amount of data loss.

RDG SCH8 AUDIT - NRPR
REPORT

2017	Unallocated Residual	R % of T	Unall % of R	Unall % of T
Arriva Trains Wales		8.9%	3.4%	0.301%
C2C		0.3%	2.2%	0.006%
Caledonian Sleeper		1.2%	8.6%	0.101%
Chiltern		1.0%	5.1%	0.051%
CrossCountry		11.2%	2.8%	0.312%
Crossrail		0.1%	2.8%	0.003%
East Midlands Trains		3.9%	4.3%	0.170%
Grand Central		0.3%	2.1%	0.007%
Great Western 1		4.4%	3.8%	0.167%
Great Western 2		2.2%	7.1%	0.153%
Great Western 3		4.7%	4.9%	0.230%
Greater Anglia		2.4%	5.1%	0.120%
Hull Trains		0.2%	0.6%	0.001%
Island Line		-0.1%	3.3%	-0.002%
London Midland		3.1%	3.8%	0.118%
London Overground		2.2%	2.1%	0.047%
Merseyrail		0.9%	6.4%	0.056%
Northern 1		3.2%	6.4%	0.204%
Northern 2		3.6%	7.1%	0.252%
Northern 3		4.2%	5.0%	0.214%
Northern 4		5.0%	4.8%	0.243%
Scotrail 1		1.6%	7.8%	0.125%
Scotrail 2		3.3%	3.4%	0.112%
South Western 1		1.7%	-0.1%	-0.001%
South Western 2		2.3%	3.8%	0.088%
South Western 3		3.0%	2.6%	0.079%
Transpennine		4.4%	4.9%	0.217%
VT East Coast		3.8%	3.3%	0.124%
VT West Coast 1		0.3%	1.2%	0.004%
VT West Coast 2		6.0%	3.3%	0.197%

FIGURE 6 RESIDUAL FLOWS LOSSES

Figure 6 shows the losses at this stage. The actual loss figures are obscured. The other three columns in this figure are:

- R % of T: proportion of the TOC total revenue represented by the residual flows.
- Unall % of R: proportion of the residual flows revenue unallocated in processing
- Loss % of T: proportion of the TOC total revenue unallocated in residual flow processing.

We did not see why any revenue should be unallocated at this stage so did some further investigation. We found one very minor issue (https://github.com/VivacityRail/RDG_PR18_Ph3/issues/12_severity_4).

We decided that the level of lost revenue was not high enough to warrant further work, particularly as the mismatch would be corrected later, in the NRPR spreadsheet, as revenues are scaled back to overall Lennon totals.

Appendix D Decision Tables and Test Cases

D.1 INTRODUCTION

The only sound way to test the functionality of a spreadsheet model is to prepare test cases and expected results and verify that the model's results match the expected values in all cases.

Decision tables are valuable in this endeavour because they make it easy to consider all the possible combinations of inputs. Using them to generate test cases means that complete coverage of the test cases can be guaranteed.

D.2 DECISION TABLES

The figures in this section show the decision tables used to represent all the possible logical conditions that flows could have.

A decision table consists of a number of questions or conditions, each of which could have a small number of answers - typically Yes or No, or Full / Reduced/ Season; and a set of possible results or actions. All the feasible combinations of conditions are set out, and actions defined for them.

Geography	LL	LR	LS	RR	SR	SS
Calculation	B	A	B	A	A	A

FIGURE 7 DECISION TABLE 1 - WHICH CALCULATION TYPE TO USE

RDG SCH8 AUDIT - NRPR
 REPORT

Geography	LL	LL	LL	LS	LS	LS
Ticket Type	F	R	S	F	R	S
Semi-elasticity						
-0.1133	y					
-0.0645		y				
-0.0437			y			
-0.0205				y		
-0.0305					y	
-0.0210						y

FIGURE 8 DECISION TABLE 2 - SEMI-ELASTICITY FOR L/SE MRE CALCULATION

Geography	AA	LR	LR	LR	LR	RR	RR	RR	RR	SR	SR	SR	SR	SS	SS	SS	SS
Dist Band	-	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
GeogDistBand	AA-	LR1	LR2	LR3	LR4	RR1	RR2	RR3	RR4	SR1	SR2	SR3	SR4	SS1	SS2	SS3	SS4
PDFH Dist Group (Local / Interurban)																	
Local (<20)						Y				Y				Y			
Interurban (>20)			Y	Y	Y		Y	Y	Y		Y	Y	Y		Y	Y	Y
Journey Purpose Dist Group																	
S (<25)			Y			Y	Y			Y	Y			Y	Y		
M (25-100)				Y				Y				Y				Y	
B (>100)					Y				Y				Y				Y
Airport	Y																
Sector																	
Long distance to/from London			Y	Y	Y												
Non-London (more than 20 miles)						Y	Y	Y		Y	Y	Y					
Non-London (less than 20 miles)						Y				Y							
South East Outer Suburban														Y	Y	Y	Y
GJT Elasticity (use sector)																	
-1.10						Y				Y							
-1.20							Y	Y	Y		Y	Y	Y				
-1.25									Y					Y	Y	Y	Y
-1.35			Y	Y	Y												
-1.50	Y																

FIGURE 9 DECISION TABLE 3 - GEOGRAPHY AND DISTANCE

This table defines the PDFH sector and the GJT elasticity to use, based on the Geography coding and distance banding (the bands are defined separately as 1-20, 20-25, 25-50, 50-100 and 100+ miles).

Note the grey column for case LR / 1 - a short-distance journey from London to Rest of UK. The actions for this are not defined in PDFH. This appears to be moot as no actual flows can meet this case.

RDG SCH8 AUDIT - NRPR
REPORT

Dist / Geog table	AA	LRB	LRM																
Ticket Type	-	F	F	F	R	R	R	S	S	S	F	F	F	R	R	R	S	S	S
JP	-	B	C	L	B	C	L	B	C	L	B	C	L	B	C	L	B	C	L
Band 1?	-	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Airport	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Sector																			
Long distance to/from London		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Non-London (more than 20 miles)																			
Non-London (less than 20 miles)																			
South East Outer Suburban																			
Delay Multiplier																			
2.3																			
2.5			Y	Y		Y		Y		Y		Y		Y		Y		Y	
3		Y		Y	Y		Y	Y		Y	Y		Y	Y		Y	Y		Y
3.4																			
3.9																			
6	Y																		

FIGURE 10 DECISION TABLE 4 - DELAY MULTIPLIER (PART OF TABLE)

D.3 TEST CASES - FLOW DATA

D.3.1 DECISION TABLE CASES

Once the decision tables have been set up, dummy flow data can be defined for each case in each table. Spreadsheet formulae are used to build the flow data entries from the decision table values.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Case	Lookupcode	Year	Origin	Destination	Geography	Distance Band	Service Code	Service Code Description	GJT Full	GJT Redu	GJT Seat	GJT Band	Distance	Revenue Full	
1	Table 1 LL B	2016	O LL B	D LL B	London to/from London	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
2	Table 1 LR A	2016	O LR A	D LR A	London to/from Rest of Country	25-100	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
3	Table 1 LS B	2016	O LS B	D LS B	London to/from South East	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
4	Table 1 RR A	2016	O RR A	D RR A	Rest of Country to/from Rest of Country	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
5	Table 1 SR A	2016	O SR A	D SR A	South East to/from Rest of Country	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
6	Table 1 SS A	2016	O SS A	D SS A	South East to/from South East	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
7	Table 2 LF F	2016	O LF F	D LF F	London to/from London	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
8	Table 2 LL R	2016	O LL R	D LL R	London to/from London	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
9	Table 2 LS S	2016	O LS S	D LS S	London to/from London	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
10	Table 2 LS F	2016	O LS F	D LS F	London to/from South East	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
11	Table 2 LS R	2016	O LS R	D LS R	London to/from South East	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
12	Table 2 LS S	2016	O LS S	D LS S	London to/from South East	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
13	Not a test case	2016	O AA 1	D AA 1	London to/from South East	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	0	0	
14	Doesn't happen	2016	O LR 1	D LR 1	London to/from Rest of Country	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
15	Table 3 LR 2	2016	O LR 2	D LR 2	London to/from Rest of Country	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
16	Table 3 LR 3	2016	O LR 3	D LR 3	London to/from Rest of Country	25-100	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
17	Table 3 LR 4	2016	O LR 4	D LR 4	London to/from Rest of Country	100+	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
18	Table 3 RR 1	2016	O RR 1	D RR 1	Rest of Country to/from Rest of Country	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
19	Table 3 RR 2	2016	O RR 2	D RR 2	Rest of Country to/from Rest of Country	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
20	Table 3 RR 3	2016	O RR 3	D RR 3	Rest of Country to/from Rest of Country	25-100	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
21	Table 3 RR 4	2016	O RR 4	D RR 4	Rest of Country to/from Rest of Country	100+	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
22	Table 3 SR 1	2016	O SR 1	D SR 1	South East to/from Rest of Country	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
23	Table 3 SR 2	2016	O SR 2	D SR 2	South East to/from Rest of Country	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
24	Table 3 SR 3	2016	O SR 3	D SR 3	South East to/from Rest of Country	25-100	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
25	Table 3 SR 4	2016	O SR 4	D SR 4	South East to/from Rest of Country	100+	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
26	Table 3 SS 1	2016	O SS 1	D SS 1	South East to/from South East	0-20	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
27	Table 3 SS 2	2016	O SS 2	D SS 2	South East to/from South East	20-25	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
28	Table 3 SS 3	2016	O SS 3	D SS 3	South East to/from South East	25-100	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	
29	Table 3 SS 4	2016	O SS 4	D SS 4	South East to/from South East	100+	2100	MARYLBONE-HARROW-AYLESBURY	20	30	40	Not Appl	100	3,200	

FIGURE 11 TEST FLOW DATA - SAMPLE

Figure 11 shows a small sample of the dummy flow data generated. Some points to note:

- Columns in the data such as Lookup Code, Origin and Destination are used to identify the test case by its decision table lookup key. This

means that any test failure can be pinned down to the specific set of conditions that caused it.

- Other data items are kept constant as far as possible between test cases. This means that the only thing causing variations in the results is the decision table conditions.

D.3.2 OTHER TEST CASES

Besides the decision table cases, other test cases were manually added to check for particular situations:

- Identification of airport flows
- Handling of residual flows, which have different columns populated in the data set
- Flow characteristics to check the lookup tables could cover all cases - e.g. looking up from the first and the last row of each lookup table to verify that spreadsheet ranges had been correctly stretched over the lookup tables.
- Extreme numerical values of revenue, GJT and distance.
- Very large data set that completely fills the allocated data input range (160000 rows, for flow data), to verify that all rows are picked up.
- Pathological flows that exhibited a variety of data problems such as missing or invalid GJT values or distances. (These were handled as a separate exercise, as including them in the normal set of test cases should cause the spreadsheet calculations to fail and so produce invalid results).

D.4 THE TEST HARNESS

A mock-up of the NRPR flow MRE calculation was put together to calculate expected results for Peak and Off-Peak MRE for the test flows. This used a data table and table formulae to calculate the MREs in a single row of formulae. Table formulae are helpful because they reference cells by their column name: this greatly improves readability of the formulae and thus the ability to debug and check the calculations. They also automatically replicate themselves over the whole column.

Appendix E Results of Detailed Tests

The test data set of 234 flows, covering all the decision table cases and a further 68 rows of manually-prepared cases, was run through the NRPR model and the test harness.

NRPR Model vs Test Harness										
2017 data		NRPR Model				Difference		Precision	Check	
Flow	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak OK	Off-Peak OK		
1	58.76	536.14	58.76	536.14	0.00	0.00	TRUE	TRUE		
2	127.50	1,034.85	127.50	1,034.86	0.00	0.00	TRUE	TRUE		
3	3.60	212.90	3.60	212.90	0.00	0.00	TRUE	TRUE		
4	102.43	1,201.97	102.43	1,201.97	0.00	0.00	TRUE	TRUE		
5	96.25	1,215.95	96.25	1,215.95	0.00	0.00	TRUE	TRUE		
6	89.99	797.51	89.99	797.51	0.00	0.00	TRUE	TRUE		
7	58.76	536.14	58.76	536.14	0.00	0.00	TRUE	TRUE		
8	58.76	536.14	58.76	536.14	0.00	0.00	TRUE	TRUE		
9	58.76	536.14	58.76	536.14	0.00	0.00	TRUE	TRUE		
10	3.60	212.90	3.60	212.90	0.00	0.00	TRUE	TRUE		
11	3.60	212.90	3.60	212.90	0.00	0.00	TRUE	TRUE		
12	3.60	212.90	3.60	212.90	0.00	0.00	TRUE	TRUE		
13	0.00	0.00	0.00	0.00	0.00	0.00	TRUE	TRUE		
14	145.66	1,040.99	145.66	1,040.99	0.00	0.00	TRUE	TRUE		
15	145.66	1,040.99	145.66	1,040.99	0.00	0.00	TRUE	TRUE		
16	127.50	1,034.85	127.50	1,034.86	0.00	0.00	TRUE	TRUE		
17	110.84	1,029.23	110.84	1,029.23	0.00	0.00	TRUE	TRUE		
18	64.34	797.84	64.34	797.84	0.00	0.00	TRUE	TRUE		
19	102.43	1,201.97	102.43	1,201.97	0.00	0.00	TRUE	TRUE		
20	81.58	1,248.62	81.58	1,248.62	0.00	0.00	TRUE	TRUE		
21	59.56	1,297.64	59.56	1,297.64	0.00	0.00	TRUE	TRUE		
22	60.58	811.61	60.58	811.61	0.00	0.00	TRUE	TRUE		
23	96.25	1,215.95	96.25	1,215.95	0.00	0.00	TRUE	TRUE		
24	140.24	1,117.96	140.24	1,117.96	0.00	0.00	TRUE	TRUE		

FIGURE 12 MRE CALCULATION TEST RESULTS - TOP ROWS

Figure 12 shows the first few rows of the results comparison. Rows are marked as OK if the results are the same to within £0.0001. In the first run of the test shown here, 8 tests gave different answers. These are listed in Figure 13. As explained in the text, these appear to be caused by errors in the test harness rather than the model.

RDG SCH8 AUDIT - NRPR
 REPORT

NRPR Model vs Test Harness									
							Precision	0.0001	
2017 data	Test Harness		NRPR Model			Difference		Check	
Fl	Pε	Off-Pε	Pε	Off-Pε	Pε	Off-Pε	Peak OK	Off-Peak	
167	14.69	134.03	225.36	1,948.37	210.67	1814.33	FALSE	FALSE	
169	0.18	10.65	39.51	2,536.32	39.33	2525.67	FALSE	FALSE	
173	2.94	26.81	269.78	2,324.96	266.85	2298.15	FALSE	FALSE	
175	2.70	159.68	13.05	824.33	10.35	664.65	FALSE	FALSE	
187	2.71	23.38	0.04	0.51	-2.68	-22.88	FALSE	FALSE	
192	89.99	797.51	0.00	887.50	-89.99	89.99	FALSE	FALSE	
208	102.43	807.05	1.28	17.04	-101.15	-790.01	FALSE	FALSE	
209	44.16	377.84	1.28	17.04	-42.88	-360.81	FALSE	FALSE	

FIGURE 13 TEST RESULTS - MISMATCHES