steer davies gleave

Origin-Destination Matrix 2013/14

Summary Report

Report
December 2014

Office of Rail Regulation

22670601

steer davies gleave

Origin-Destination Matrix 2013/14

Summary Report

Report
December 2014

Office of Rail Regulation

22670601

Prepared by:

Steer Davies Gleave
28-32 Upper Ground
London SE1 9PD

+44 (0)20 7910 5000
www.steerdaviesgleave.com

Prepared for:

Office of Rail Regulation
One Kemble Street
London WC2B 4AN

Contents

Tables

Appendices

Executive Summary

Introduction

1. The Origin Destination Matrix (ODM) forms a vital part of the Office for Rail Regulation (ORR)'s information about how passengers travel on the railways in England, Wales and Scotland. The ODM gives information for revenue and journeys, by ticket type, for each rail flow across the country, i.e. each combination of origin station, destination station and ticket route code.

2. This report is provided with the ODM file, and gives guidance on the methodology that has been followed during the process of creating the dataset for financial year 2013/14 (1st April 2013 to 31st March 2014).

3. The ODM shows the numbers of journeys made, and resulting ticket revenue and passenger miles, for each flow (pair of origin and destination stations) in Great Britain. Where tickets are offered via different routes, the data is also broken down into those routes. It is used as the source for the ORR's regional usage profiles in statistical release publication. If further analysis is needed ORR may be able to respond to such requests.

4. Tickets are offered between every pair of stations in Great Britain, though not all combinations register a sale in any particular year. For each pair of stations, journeys and revenue figures are split between four different ticket types and between standard and first classes.

5. While LENNON is the major source of data for the ODM, it is augmented by a range of additional data sources to provide a more complete representation of travel on the national rail network. Since 2008/09, this has included estimates of journeys and revenue made in major urban areas on PTE[1] sponsored tickets which were previously excluded due to issues of distributing passenger journeys to flows. In subsequent years a number of improvements have been made to the methodology used to represent journeys associated with PTE-sponsored tickets. Notwithstanding the improvements made to represent passenger journeys in the ODM, there are limitations on the data which users should be aware of and which are detailed in this report.

Methodological Development

6. Consistency with past datasets is important to enable comparisons to be made over time. However, stakeholders have indicated that they are keen to see improvements, even where this leads to inconsistency with historic data, provided changes are clearly explained.

[1] Passenger Transport Executives (PTEs) are local government bodies which are responsible for public transport within large urban areas. They are accountable to Integrated Transport Authorities (ITAs) which were formerly known as Passenger Transport Authorities (PTAs) prior to 2008 and the Local Government Act 2008. There are five PTEs in England, for each of the metropolitan counties (Merseyside, South Yorkshire, Tyne and Wear, West Midlands and West Yorkshire) with the former Greater Manchester Passenger Transport Executive being replaced by Transport for Greater Manchester from April 2011. In Scotland the Strathclyde Partnership for Transport is the equivalent body covering the region of Strathclyde. For convenience in this report we continue to refer to these areas as PTEs.

7. In the 2013/14 dataset a number of changes have been made to improve the representation of journeys on PTE-sponsored tickets in the following PTEs:

   - Merseyside;
   - Strathclyde; and
   - South Yorkshire

8. **It is important to note that differences between this year and previous years' figures on these flows need to be considered in light of design changes to the methodology which affect the level and distribution of demand across flows.**

9. As in 2011/12 and 2012/13, the generic methodology for separating out demand at group stations to their component stations has not been followed for Manchester BR, Wigan BR and Warrington BR[2].

Limitations of the data

10. The use of ticket sales data, LENNON, as the primary source for the development of the ODM, as described in this report, is the best approach available. In particular its national coverage makes it suitable as a basis for the production of official statistics such as those reported by the ORR.

11. However, this data does have weaknesses and, although some of these are catered for in the methodology, the user should be aware of these acknowledged limitations and bear these in mind when using the data. The key limitations are outlined in Chapter 1 with more extensive discussion of some aspects of the limitations of the dataset included in Appendix B.

[2] For Warrington BR and Wigan BR, the 2010/11 split of journeys has been maintained at a flow and route code level. For Manchester BR the 2010/11 split has been maintained at the station level.

1 Introduction

1.1 Steer Davies Gleave was appointed by the Office of Rail Regulation (ORR) to produce the ODM for 2013/14, continuing the historic series that dates back to 1997/98. This report accompanies the ODM for 2013/14 and provides details of the process and outputs used to produce the dataset on behalf of the ORR.

1.2 The methodology adopted by Steer Davies Gleave in the production of the ODM is generally consistent with that adopted by DeltaRail in the production of the ODM prior to 2011/12. As part of our work we undertook a Methodological Review in 2012 of the data and processes used to generate the ODM and identified a number of areas for improvement in the data set. A number of these were implemented in the 2011/12 and 2012/13 datasets and a further set of changes has been implemented in the 2013/14 dataset (see Chapter 4).

Use of statistics sourced from the ODM

1.3 When using statistics based on the ODM (e.g. station usage data also published by the ORR) it is important to be aware of:

i. Improvements made to the dataset over time which can impact consistency between years;
ii. Limitations of the data and specifically factors e.g. some ticket sales not being included, that may mean that demand on particular flows or stations is underestimated; and
iii. Factors which can affect reporting of entries and exits.

Improvements to the dataset

1.4 Improvements to the dataset in 2013/14 are set out in Chapter 4 and mainly relate to the updated methodology for PTE-sponsored tickets in the Merseyside, South Yorkshire and Strathclyde PTEs. A summary of improvements made over recent years are further detailed in Appendix A. The ORR continues to work with stakeholders and its own consultants to improve the robustness of the dataset by implementing methodological changes that demonstrate value and address acknowledged issues.

Limitations of the data

1.5 In the absence of a completely gated system that allows a complete recording of flows through stations or comprehensive and robust count data, the use of ticket sales data, LENNON, as the primary source of the ODM is the best approach available. In particular, its national coverage makes it suitable as a basis for the production of national statistics such as those reported by the ORR. However, this data does have weaknesses when utilised for this purpose and, although some of these are catered for in the methodology, the user should be aware of these acknowledged limitations. The key limitations are outlined below. More extensive discussion of some aspects of the limitations of the dataset is included in Appendix B.

- **Non-Point to point tickets** - An overarching issue is the inherent difficulty and uncertainty associated with estimating the number of journeys associated with many rail products which do not simply represent point to point single or return journeys and furthermore

the distribution of those journeys. This is a particular issue for the London Travelcard Area and PTE areas;

- **Concessionary travel** – Most PTEs subsidise some form of free travel for passengers over a certain age and those with disabilities. This creates a substantial additional element of demand which is very difficult to include in the ODM as information on the level and distribution of journeys associated with these free travel products is not recorded. The current approach to this in the ODM is to include this demand where data has been made available by PTEs which would generally be estimates as a result of surveys. In addition, since 2012/13 an estimate of Freedom Pass journeys in the London Travelcard Area has been included;
- **Non-LENNON sales** - A significant proportion of sales is either not passed directly through LENNON (sold at non-railway sales points) or is included in LENNON in a format which requires additional processing and assumptions i.e. is not associated with a station to station flow;
- **Group stations** – Many products to major destinations are sold with the origin or destination as a group of stations (e.g. London Terminals, Manchester BR stations). Current industry data does not distinguish between the component stations and therefore a split between these stations has to be estimated during the production of the ODM; and
- **Ticketless travel** – Journeys associated with ticketless travel are not included in the datasets but as with journeys made on other products excluded from the datasets, some journeys would be observed in passenger counts. This is likely to be an issue on some flows and in some areas where ticketless travel is significant. As more stations have become gated over time and TOCs focus on revenue protection activities this is likely to be less of an issue than in the past in contributing to a shortfall in journeys. Finally, there is a strong argument that it is inappropriate to include ticketless travel in the ODM as its purpose is to record bona-fide journeys on the rail network and inclusion of ticketless travel could distort business cases for new investment where these are reliant on data from the ODM.

1.6 It is important to remember that in aggregate the underlying data, from LENNON, is a rich and comprehensive data source and, importantly, covers the entirety of Great Britain. The issue is that when using the data source to construct the ODM the data is being pushed significantly beyond what it was originally designed for which was primarily to report and allocate revenues across train operators.

2 Matrix Definition

2.1 The ODM contains revenue, journeys and passenger miles data for each flow on the network. A flow is defined as an origin station / destination station / ticket type/ route code combination. Since this dataset is designed to show passenger journeys made, rather than "producer-attractor" figures, journeys have been split equally into the two directions of travel. The fields included in the ODM are:

Table 2.1 ODM fields

Field	Description
Mode	This variable is used to categorize the source of the passenger journey data. Refer to Table 2.2 below.
Origin (NLC, name)	Based on ticket origin, assumed to be where passenger starts his/her journey.
Destination (NLC, name)	Based on ticket destination, assumed to be where passenger ends his/her journey.
District, County, Region and NUTS2 Region & Code for Origin	Origin's geographical location.
District, County, Region and NUTS2 Region & Code for Destination	Destination's geographical location.
Route Code and Description	Route code and description on ticket as recorded by LENNON.
Dist	Distance in miles between origin and destination derived from LENNON.
Revenue	Revenue for each flow split into the eight ticket types. It is also summarised into the four main categories (Full, Reduced Excluding Advance, Advance and Seasons) and a Reduced category (Reduced plus Advance) and summarised in total.
Journeys	Journeys for each flow split into the eight ticket types. Journeys are also summarised into the four main categories (Full, Reduced Excluding Advance, Advance and Seasons) and a Reduced category (Reduced plus Advance) and summarised in total.
Passenger Miles	Miles the passengers travelled - effectively journeys multiplied by a station to station distance derived from LENNON.
Group Station (NLC, name) for Origin	If the origin is part of a Group Station, the NLC and name is provided, otherwise this field is blank.
Group Station (NLC, name) for Destination	If the destination is part of a Group Station, the NLC and name is provided, otherwise this field is blank.
Flag	Flag = 0 (no problem), 1 (flow has failed a check), or 2 (flow has failed a check and may be significant).

Table 2.2 Mode definitions

Mode	Description
NR Sold Non-Tcard	Sold by National Rail, point to point
TfL Sold Tcard	London Travelcards sold by Transport for London
NR Sold Tcard	London Travelcards sold by National Rail
PTE Sold	Sales of PTE-sponsored tickets
Airline Sold	Ticket sales for routes serving Airports, where tickets do not go through LENNON
Other	A small number of Rangers and Rovers and non-LENNON season products

3 Methodological Overview

MOIRA2 Demand Matrix – Base Data

Overview

3.1 The ODM is derived primarily from the MOIRA2 Demand Matrix. The MOIRA2 demand matrix is sourced from MOIRA2 which is the rail industry's principal planning tool and includes a comprehensive representation of travel on the national rail network. The base data for the MOIRA2 demand matrix is LENNON ticket sales, with the addition of "infills" for London Travelcards, airport links and multi-modal and zonal products sponsored by Passenger Transport Executives (PTEs).

Underlying Base Data - LENNON

3.2 The underlying matrix of ticket sales and associated journeys and revenue used in MOIRA2 is derived from LENNON. It is based on an extract from LENNON, produced by Atos, of total sales revenue and journeys for the year, broken down by flow (origin and destination National Location Code (NLC)), route code and by product type (CTOT). There are known omissions in this data in respect of Transport for London (TfL) and PTE sponsored tickets, and non-National Rail tickets on some airport services. As a result there needs to be a "matrix infilling" exercise undertaken to estimate a more complete origin-destination matrix and include the associated journeys and revenue that do not appear in the underlying matrix.

3.3 There are three main cases:

- Tickets with non-geographical destinations, e.g. zonal products, Rovers;
- Tickets sold at some non-National Rail (RSP: Retail Settlement Plan) outlets, e.g. newsagents; and
- Tickets which do not appear in LENNON at all. This includes some Train Operating Company (TOC) tickets on airport flows, and tickets for TOCs which fall outside the Rail Settlement Plan.

3.4 Certain tickets with destination codes that are not National Rail stations are included in the MOIRA2 demand matrices, being mapped to the corresponding rail station. These Rail Links usually include a third party element, such as to a bus zone, or tourist attraction. The MOIRA2 demand matrix includes the journeys and the net revenue associated with such tickets.

3.5 Data excluded from the MOIRA2 demand matrix is set out in Appendix B.

Net Revenue

3.6 The MOIRA2 demand matrix contains Net Revenue based on the "Net Revenue" field in LENNON. Travelcard revenue in MOIRA2 is Net (rather than Gross) i.e. excludes revenue paid by TOCs to TfL for travel on the London Underground and on buses. Similarly, PTE revenue is net i.e. for multi-modal tickets only revenue associated with travel on national rail services is included.

Ticket Type Definitions

3.7 Within the base demand matrices, journeys and revenue have been sub-divided into the following four ticket types, each of which is further split by First & Standard Class:

- Full: all walk-up undiscounted single or return tickets, whether or not issued with a status discount (child, railcard etc);
- Reduced: all walk-up discounted single or return tickets, whether or not issued with a status discount (child, railcard etc);
- Advance: all advance-purchase tickets;
- Seasons: all multi-use tickets.

Infills for London Travelcards, Major Urban Areas (PTE) & Airports

3.8 Infills are included within the MOIRA2 demand matrix to add in the missing journeys and revenue identified in para 3.3 in three key areas:

- **Within London Travelcard area -** Whilst the underlying matrix includes an estimate of journeys made on Day Travelcards / Travelcard seasons purchased at National Rail stations, it does not include a significant number of national rail trips made using Travelcards purchased at Tube stations, travel shops and newsagents.
- **Within Passenger Transport Executive (PTE) areas -** The underlying matrix excludes virtually all rail trips made on PTE-sponsored tickets, which are usually zonal and often multimodal.
- **Trips to/from Airports -** The underlying matrix includes many trips to/from airports, but excludes all Heathrow Express journeys, and some tickets sold for Gatwick Express, Stansted Express and other airport operators.

3.9 There are also other ticket sales which are not included in the MOIRA2 demand matrix, but these are generally much less significant. It should also be noted that journeys with no associated ticket sales such as staff travel, and particularly fare evaders, are not included in the MOIRA2 demand matrix and therefore are not included in the ODM either.

3.10 The most significant “infills” are for the London Travelcard area (sales made by Transport for London (TfL)), and for PTEs, since in both cases a substantial proportion of the rail journeys made use multimodal travelcard type of tickets.

3.11 The third infill, for Airports, estimates the significant number of rail journeys on Gatwick and Stansted Express, made on tickets sold outside of the RSP system i.e. not sold by National Rail outlets. Journeys on Heathrow Express are excluded from the MOIRA2 demand matrix.

PTE Infill

3.12 For the production of the ODM the revenue and journeys associated with the MOIRA2 PTE Infill are removed and replaced with a separate estimate.

3.13 With the initial version of MOIRA2 an improved representation of PTE demand was included in the base demand matrix based on work undertaken by Steer Davies Gleave for the year 2008/09. This included journeys from tickets sold at non-railway sales points and an estimated distribution of journeys largely based on the distribution of point to point tickets sold in PTE areas.

3.14 Subsequent versions of the MOIRA2 demand matrix have included a PTE infill but the journeys are now based directly on LENNON data and are therefore not consistent with the 2008/09 infill.

3.15 To maintain consistency with previous ORR statistics the PTE infill contained in the ODM has therefore been based on the 2008/09 MOIRA2 PTE infill (as described in para 3.13) grown by growth rates derived from ORR’s National Statistics data.

3.16 Up until 2010/11 the application of growth was carried out at a highly aggregate level based on growth seen for 'franchised regional operators' as reported in ORR's National Statistics data. In the construction of the 2011/12 dataset a more disaggregate set of growth rates was applied at the PTE level based on LENNON data. In addition, a completely new infill was included for the West Midlands PTE (Centro) based on an infill constructed for the Passenger Demand Forecasting Council (PDFC) by Steer Davies Gleave.

3.17 In 2012/13, new infills were included for Greater Manchester and West Yorkshire PTEs. These have also been included in 2013/14.

3.18 Additionally in 2013/14, new infills have also been included for the Merseyside, South Yorkshire and Strathclyde PTE areas and these are detailed in Chapter 4. In summary, as a result of these methodological enhancements over the last two years users should be cautious in the comparisons they make over time for station usage levels in these areas.

Table 3.1 Summary status of PTE infills methodology

PTE	Status
Greater Manchester	Updated infill methodology adopted for 2012/13 and 2013/14
Merseyside	Updated infill methodology adopted for 2013/14
South Yorkshire	Updated infill methodology adopted for 2013/14
Strathclyde	Updated infill methodology adopted for 2013/14
Tyne & Wear	Original 2008/09 methodology maintained.
West Midlands	Updated infill methodology adopted for 2011/12 through to 2013/14.
West Yorkshire	Updated infill methodology adopted for 2012/13 and 2013/14

4 Methodological Changes in 2013/14

Introduction

4.1 Consistency with past datasets is important to enable comparisons to be made over time. However, stakeholders have indicated that they are keen to see improvements, even where this reduces consistency with historic data, provided any changes are clearly explained.

4.2 In the 2013/14 dataset a number of changes have been made to improve the dataset and these are explained in the rest of this chapter, together with some quantification of their impact.

South Yorkshire PTE Infill

4.3 Building on the inclusion in the 2012/13 dataset of an improved infill for the West Yorkshire (WYPTE) and Greater Manchester (GMPTE/TfGM) PTE areas, an improved infill for the South Yorkshire (SYPTE) PTE area has been included in the 2013/14 dataset. This was produced using a process derived to construct infill demand for the Rail in the North demand and revenue model produced by Mott MacDonald and MVA for the Rail in the North (RiN) consortium and was supplied by Mott MacDonald. This is consistent with the methodology underlying the improved West Yorkshire (WYPTE) and Greater Manchester (GMPTE/TfGM) infills. At the total PTE level the impact of the new infill has been to reduce demand by approximately 0.7m journeys[3].

4.4 It is important to note that this reduction in journeys is not necessarily indicative of any underlying reduction in demand in the South Yorkshire PTE area but rather the result of the methodological changes implemented in this year's data.

Merseyside PTE Infill

4.5 Currently the infill for the Merseyside area is derived from the generic PTE infill produced as part of the MOIRA2 Replacement project which was based on a 2008/09 base year. To produce updated estimates in succeeding years, the distribution of demand in the infill matrix has been maintained and the total volume of demand grown, initially by the journey growth shown by the Regional Sector in the ORR's rail usage data and, since 2011/12, by the growth in journeys (from LENNON) on service codes associated with the Merseyside area.

4.6 Since 2008/09 there have been a number of developments which mean that the 2008/09 distribution is inappropriate. Of particular importance has been a movement away from RSP products to PTE products on some routes on the edges of the Merseytravel area (e.g Town Green, Aughton Park and Ormskirk on the Northern line) which means that the existing distribution underestimates demand in these areas.

4.7 Recognising the deficiencies of the existing infill, a new infill has been produced by Mott MacDonald building on the PTE infill in the Liverpool City Region Model (LCRM) produced for

[3] As all the new Mott MacDonald infills were incorporated into the ODM at the same time, it is not possible to definitively isolate each infill. For the purposes of this exercise, stations within the Yorkshire and Humber Government Office Region were considered to be those affected by the new SYPTE infill.

Merseytravel. Unlike the other PTE infills, journeys in the Merseyside infill have been scaled to count data at an aggregate level across all affected stations where complete counts are available to ensure a robust match with 'reality'. This is possible since count data in the Merseyside area is more extensive and comprehensive across stations than in other areas.

4.8 The inclusion of the new infill increases journeys by approximately 5.4m (5.1% of total North West journeys)[4].

4.9 It is important to note that this increase in journeys is not necessarily indicative of any underlying increase in demand in the Merseyside PTE area but rather the result of the methodological changes implemented in this year's data.

Strathclyde Passenger Transport (SPT) infill

4.10 A more sophisticated infill has been developed by Mott MacDonald to capture demand in the Strathclyde area on a number of SPT products, namely:

- Zonecard;
- Roundabout; and
- Daytripper

4.11 Total sales data for these tickets has been obtained from a combination of LENNON data and off rail sales figures from SPT. The number of journeys on each ticket type has been established by applying appropriate tip rate proxies for each type. The data has been distributed using Zonecard forum travel diary data and LENNON station-station reduced ticket proportions to produce an estimate of station-to-station movements. The new infill results in a drop of approximately 2.2m journeys (2.5% of Scotland's total journeys)[5].

4.12 It is important to note that this reduction in journeys is not necessarily indicative of any underlying reduction in demand in the Strathclyde area but rather the result of the methodological changes implemented in this year's data.

Other methodological variations

4.13 As for 2011/12 and 2012/13 the generic methodology for separating out group stations has not been followed for Manchester BR, Wigan BR and Warrington BR. For Warrington BR and Wigan BR we have maintained the same split of journeys between the respective stations as seen in 2010/11 at a flow and route code level. For Manchester BR the split has been maintained at the station level.

[4] As all the new Mott MacDonald infills were incorporated into the ODM at the same time, it is not possible to definitively isolate each infill. For the purposes of this exercise, stations within the North West Government Office Region were considered to be those affected by the new Merseyside infill.

[5] As all the new Mott MacDonald infills were incorporated into the ODM at the same time, it is not possible to definitively isolate each infill. For the purposes of this exercise, stations within the Scotland Government Office Region were considered to be those affected by the new SPT infill.

steer davies gleave

Appendices

A Appendix – Historical Methodological Changes

Historical Methodological Changes

A.1 In the five years prior to the 2011/12 dataset a number of improvements were made to the ODM and Station Usage methodology which are described in this section. This includes the inclusion since 2009/10 of Oyster PAYG data in the ODM which represented a significant improvement to the estimates for rail travel across London (these improvements are described in the first section of this Appendix).

A.2 The 2011/12 dataset included a raft of improvements reflecting recommendations in the Methodological Review undertaken by Steer Davies Gleave in 2012 and details of these improvements are included as a separate section in this Appendix along with further improvements included in the 2012/13 dataset.

Historical methodology changes prior to 2011/12

A.3 Between 2006/07 and 2008/09 the accuracy and usefulness of the ODM was improved by applying new procedures on the way journeys with unknown origin and/or destination have been treated, and by including journeys that were previously excluded from the file or did not appear in the LENNON sales data. In summary, the main changes were:

- Adding in previously missing journeys, e.g. TfL sold Travelcards, and some airport link tickets -this is undertaken in the production of the MOIRA2 demand matrix.
- Rail Links such as PlusBus and Attractions. The rail element of these ticket sales is now included - this is undertaken in the production of the MOIRA2 demand matrix.
- Estimating the split of records for station groups, including London BR, into the constituent individual stations. This methodology was further refined for those groups with no ticket office at one or more stations within the group - this processing is undertaken in the ODM,
- Via the integration with the process that creates the MOIRA2 Demand Matrix, PTE ticket sales are now included, in addition to TfL sold Travelcards, and some airport link tickets – this is undertaken in the production of the MOIRA2 demand matrix.
- The method for estimating passenger journeys from ticket sales has changed. This is a result of using the MOIRA2 Demand Matrix as a starting point. The MOIRA2 Demand Matrix does not disaggregate single journeys, and so when estimating passenger journeys all ticket sales have been split equally into the two directions of travel. This will only have an impact on the ODM if there is more travel on single tickets away from a station compared to travel to the station, which is not likely to be material. Therefore in the Station Usage file, entries are the same as exits.

A.4 In 2009/10 further improvements were made:

- Adding in data for journeys undertaken by Oyster “pay-as-you-go” (PAYG) in the London area. This is undertaken within the base LENNON data, in the production of the MOIRA2 demand matrix. This applies to journeys made after 1 January 2010.
- Refinement of the methodology used to calculate journeys undertaken using PTE tickets.

A.5 When the 2010/11 dataset was constructed it emerged that the original 2008/9 figures which were given for one PTE, West Yorkshire, were not a complete record of all the rail journeys on multimodal tickets which should have been included in the PTE infill. A correction was therefore made by uplifting the West Yorkshire PTE Infill, both revenue and journeys figures, by 53% on top of the generic PTE infill growth rate. Note that within West Yorkshire PTE area, the majority of rail journeys are made on rail-only tickets, i.e. not PTE Infill tickets. Thus the overall effect of this correction was relatively small.

Oyster PAYG

A.6 Oyster 'Pay As You Go' (PAYG) was rolled out at National Rail stations in January 2010. Prior to this date Oyster PAYG was available on selected routes only and was not recorded (in LENNON) on a flow or station basis. After this date Oyster PAYG was available at all National Rail stations in the Travelcard Area and recorded by flow.

A.7 The 2009/10 data contained roughly 9 months of data prior to January 2010 and 3 months of data after, while the 2010/11 data which was wholly after January 2010 when Oyster PAYG, with data capture, had been fully implemented contains a full year of data. This lead to some very large reported growth figures for some stations within the London Travelcard (/Oyster PAYG) area. The 2010/11 figures, based on recorded use of Oyster PAYG should be accurate, but the percentage growth may be over-represented since the old figures will be largely estimates made without the benefit of Oyster records.

Methodological Improvements in 2011/12

Improved PTE Infill growth rate

A.8 With the initial version of MOIRA2 an improved representation of PTE demand was included in the base demand matrix based on work undertaken by Steer Davies Gleave for the year 2008/09. This included journeys from tickets sold at non-railway sales points and an estimated distribution of journeys largely based on the distribution of point to point tickets sold in PTE areas.

A.9 Subsequent versions of the MOIRA2 demand matrix have included a PTE infill but the journeys are now based directly on LENNON data and are therefore not consistent with the 2008/09 infill.

A.10 To maintain consistency with previous ORR statistics the PTE infill contained in the ODM was therefore based on the 2008/09 MOIRA2 PTE infill grown by growth rates derived from the ORR's National Statistics data.

A.11 Up until 2010/11 the application of growth was carried out at a highly aggregate level based on growth seen for 'franchised regional operators' as reported in ORR's National Statistics data. In the construction of the 2011/12 dataset a more disaggregate set of growth rates were applied at the PTE level based on LENNON data to improve the appropriateness of the growth rates applied and reflect geographical variations in demand growth.

Inclusion of revised West Midlands PTE (Centro) Infill

A.12 Steer Davies Gleave were commissioned in 2011 by the Passenger Demand Forecasting Council (PDFC) to construct a PTE infill matrix for the Centro area for the rail year 2010/11. The methodology followed that used for the construction of the original MOIRA2 infill but included use of additional data sources and specific adjustments for known issues such as directionality.

A.13 This infill represented a significant improvement on the infill in the ODM and therefore as part of the 2011/12 update the PDFC infill was updated to 2011/12 data and included in the ODM and hence the Station Usage dataset.

A.14 The inclusion of the Centro infill represented a significant change for stations within the Centro area and also a number of stations not in the Centro area but where Centro tickets can be purchased for travel into the Centro area. For the majority of stations the inclusion of the infill resulted in an increase in entries and exits although in a small number of instances there was a decrease. A comparison of the 2011/12 Centro infill with the 2010/11 ODM infill is included in Appendix Table A.1. This shows that the new infill added approximately 5 million journeys compared to what would have been derived had the previous methodology been used.

Appendix Table A.1 Centro area infill comparison

	2010/11 ODM infill	2010/11 infill grown to 2011/12 using previous methodology	2011/12 updated infill
Journeys (m)	15.5	16.6	21.3

New 'Other' infill layer

A.15 In some non-PTE areas there are zonal products which are not captured within the MOIRA2 demand matrix (e.g. Rover and Ranger products). Whilst volumes of travel on these tickets are relatively small, in the area of use they can be significant. Therefore, in the 2011/12 update we included journey estimates for a number of Rover and Ranger products. These were:

- St Ives Group Day Ranger
- St Ives Day Ranger
- St Ives Family Day Ranger
- Valleys Night Rider
- Cambrian Coaster Ranger

A.16 Journeys on these products were included as an 'Other' infill in the ODM, together with journeys from some non-LENNON season ticket products previously included in the airport flow infill. Journey estimates for these products were constructed using LENNON data and distributing journeys based on point of sale and the underlying reduced ticket travel distribution of the stations covered.

A.17 The total number of entries and exits arising from inclusion of these journeys was 760k. Appendix Table A.2 lists the top five stations impacted most significantly:

Appendix Table A.2 Top Five Stations Impacted By Inclusion Of The 'Other' Infill

NLC	Station Name	2010/11 entries and exits	2011/12 entries and exits	Reason
3538	St.Ives	258,530	578,214	Inclusion of St Ives branch line rover products
3542	Carbis Bay	55,334	206,736	
3537	St.Erth	120,770	202,362	
3498	Lelant Saltings	17,224	101,284	
3899	Cardiff Central	11,259,968	11,502,080	Inclusion of Valley Night Rider product

Methodological Improvements in 2012/13

Improved Greater Manchester and West Yorkshire PTE Infill

Methodology

A.18 Building on the inclusion in the 2011/12 dataset of an improved infill for the Centro area, an improved PTE infill has been included in the 2012/13 dataset for two of the remaining PTEs – West Yorkshire (WYPTE) and Greater Manchester (GMPTE/TfGM). This has been produced using a process derived to construct infill demand for the Rail in the North demand and revenue model produced by Mott MacDonald and MVA for the Rail in the North (RiN) consortium for the year 2011/12. This element of the data has been supplied by Mott MacDonald and we acknowledge the contributions of the PTEs in supplying the data to allow this methodological enhancement to be progressed.

A.19 The original Rail in the North PTE infill also covered South Yorkshire PTE tickets. However, South Yorkshire PTE was unable to release the requested data within the required timescale and therefore the current PTE infill methodology has been maintained for this PTE for 2012/13. Some data was obtained from South Yorkshire PTE in the form of numbers of boardings at main stations (such as Doncaster), with limited data available from smaller stations (such as Swinton). This data was not compatible with the methodology which was developed for the Rail in the North work, and a decision was taken not to make use of it.

Impact of change

A.20 The impact of the methodological change at the PTE level is shown in Appendix Table A.3:

Appendix Table A.3 West Yorkshire and Greater Manchester PTE infill (2012/13)

PTE	Journeys (m)	
	Old Methodology	New Methodology
West Yorkshire PTE	6.83	8.67
Greater Manchester PTE	5.05	5.10

Source: SDG Analysis of PTE infill based on a station classification into PTEs – this necessitates a simplified treatment of cross-PTE boundary flows

A.21 The new infill has a significant impact at the total level for the West Yorkshire PTE area with a 27% increase in the number of journeys on West Yorkshire PTE tickets. The impact on the total size of the GMPTE infill is much smaller but there are still significant distributional impacts across stations.

Inclusion of Freedom Pass journeys in PTE Infill

Methodology

A.22 The Tfl concessionary product the 'Freedom Pass' is included in the Oyster system. Unlike paid for Oyster products, travel on the Freedom Pass is not included in the current station usage estimates. Given the volume of rail travel on the Freedom Pass (circa 11 million journeys in 2012/13) inclusion of these journeys where possible in the station usage dataset is highly desirable.

A.23 TfL have provided the following data to enable an estimate of Freedom Pass journeys on the rail network:

- Total journeys on Freedom Pass with touch in/out at least one end of the journey at a 'NR subsystem'[6] station for each period in the 2012/13 year
- Origin and destination breakdown of Freedom pass journeys where the passenger touched in or out for period 4 of 2012/13 (July 2012), including a distinction between London Underground and National Rail services e.g. entries and exits at London Bridge National Rail and London Bridge London Undergound are recorded separately

A.24 Inclusion of the Freedom Pass journeys has been achieved through a two-stage process:

- Calculation of period 4 Freedom Pass journeys on National Rail/London Overground services by assigning each origin destination in the sample period 4 data as being either a National Rail/London Overground journey or not. This was required to exclude journeys not on the National Rail/London Overground network
- Estimation of total 2012/13 Freedom Pass journeys on National Rail/London Overground by flow by using the periodic 'NR subsystem' data to uprate the period 4 journeys

A.25 The number of Freedom Pass journeys included is necessarily a conservative estimate since it does not capture journeys where the passenger did not have to touch in or out. In addition,

[6] The NR subsystem is a set of stations which is used for recording purposes by TfL. It is composed primarily of National Rail stations but does include some joint stations (e.g. Wimbledon). As such it cannot be used to provide a completely clean estimate of total National Rail Freedom Pass journeys but the periodic data can be used to scale the detailed Period 4 data to the whole year.

the smallest flows in the period 4 dataset have not been included since it was not practical to categorise every single flow.

Impact of change

A.26 Inclusion of Freedom Pass journeys has added 10.6m journeys into the ODM matrix.

Additions to the 'Other' infill layer

Methodology

A.27 In 2011/12 a number of zonal products outside PTE areas which were not captured within the MOIRA2 demand matrix were included for the first time in the dataset as part of a new 'Other' infill layer. In the 2012/13 dataset a further five non-PTE zonal products have been included:

- Anglia Plus
- Devon Evening Ranger
- Devon Day Ranger
- Ride Cornwall
- Freedom Travel Pass (West of England product)

A.28 Journey estimates for these products have been constructed using LENNON data and distributing journeys based on point of sale and the underlying reduced[7] ticket travel distribution of the stations covered.

Impact of change

A.29 The total number of entries and exits arising from inclusion of these journeys is 1.05m

[7] With the exception of the Anglia Plus product which has both Reduced and Season variants. For the Season variants of this product the underlying Full ticket travel distribution of the stations covered was used given that the coverage of Season tickets in the base matrix was limited.

B Appendix – ODM Limitations

Limitations of the LENNON data

B.1 The LENNON database captures ticket sales for the entire national rail network from many different input machines. It is as a consequence a very large data set. With all large data sources there will always be input errors resulting in a certain amount of invalid data. Generally such errors will be small, and are more likely to occur in the journeys rather than revenue fields.

B.2 Checks are performed on the data when the MOIRA2 demand matrix is compiled, but due to the size and complexity of the dataset it is not possible to validate each and every entry.

B.3 There are a number of areas where we know that LENNON does not capture the data correctly, or instances where it is not possible to derive passenger journeys from ticket sales data. These areas are expanded upon below.

Known Problems of Data Capture

B.4 The data in LENNON from which the ODM is derived is based on ticket transactions. In order for the data to be included in the ODM it must include an origin station and a destination station. However if this is not the case then the data will automatically be excluded.

B.5 Human error at the point the ticket sale is entered into the input machines will also produce invalid data in LENNON.

Travelcards

B.6 As Travelcards are for multi-modal travel they allow the purchaser to make journeys on the rail system and on other modes. Equally, tickets purchased elsewhere on the local transport system will be valid for rail travel. Therefore LENNON gives only a partial picture of the rail travel in conurbation areas, such as: London, Birmingham, Glasgow, Leeds, Liverpool, Manchester, Newcastle and Sheffield.

B.7 The ODM contains reasonably robust estimates of journeys within London and other conurbation areas where travelcards are widely used. An infill for London Travelcards has been included in the ODM since 2006/07, and an infill for PTE tickets is included from 2008/09.

Return and Single Journey Tickets

B.8 It is possible that on certain routes the cost of a return ticket could be lower than a single ticket. This leads to the cheaper return ticket being purchased even though the passenger has no intention of making the return journey by rail. This results in two journeys being recorded instead of one.

Multiple Tickets

B.9 It is possible to buy special cheaper tickets between certain stations for example under a promotion by one of the train companies. In these cases a local ticket may be bought to gain access to a main station and a second ticket bought for the rest of the journey. This results in two journeys being recorded in the ODM and will not accurately represent the journey undertaken.

Rail Staff Passes

B.10 Prior to the privatisation of the rail network, British Rail employees and their families were eligible to various levels of free or reduced rate rail travel. When the various rail companies were converted to private companies, this benefit often continued.

B.11 If you consider the network as a whole, the effect of staff passes is unlikely to be significant. However, it may be significant on certain routes, for example on routes out of Derby due to large concentration of companies in Derby relating to British Rail both pre and post privatisation.

B.12 Ticketless Travel On every route on the network there will always be passengers who travel without purchasing a ticket. This is referred to as ticketless travel. As LENNON data is derived from ticket transactions it cannot reflect this travel.

Other Rail Systems

B.13 There are a number of rail systems in operation in the country that are not covered by LENNON. For Heathrow Express and Eurostar revenue and journeys data were not available.

Journey Factors

B.14 Ticket transactions are converted into an estimate of the number of journeys made by applying a series of ticket type journey factors. Single and return tickets unambiguously translate into one and two journeys respectively, for season tickets, the factors used represent a rough historic estimate as set out in Appendix Table B.1 overleaf.

B.15 Ticket periods of other lengths are converted to a number of journeys using a proportion of the monthly journey factor.

B.16 Therefore the journeys data in the ODM represents an assumed number of journeys made based on the ticket type sold and the above journey factors. In particular it should be noted that the journeys data has not been cross-checked against other data sources of the actual number of journeys made on the network.

B.17 These journey factors have been used within the LENNON system for a number of years at their current values. The source of the factors is unclear, and there is some indication that they were based on reasonable estimates of ticket use made in excess of fifteen years ago. It can therefore be argued that these journey factors do not provide an accurate estimate of the number of journeys that result on the rail system at present, or in any ODM.

Appendix Table B.1 Journey Factors used in LENNON

Description	Journeys Per Issue
Single Journey Ticket	1
Return Journey Ticket	2
Return Journey 2 Persons	4
3 Day Return/ 6 Single Journeys	6
4 Day Return/ 8 Single Journeys	8
5 Day Return/ 10 Single Journeys	10
6 Day Return	12
5 Day Single	5
1.5 Journeys	1.5
Weekly Ticket	10.3
10 Day Return/ 20 Single Journeys	20
2 Weekly Ticket	22
Seasons-Variable Periods	***
Monthly Ticket	45
Not Used	0
3 Monthly Tickets	135
Not Used	0
6 Monthly Tickets	270
Summary Group Codes	***
Annual Ticket	480
8 Day Ticket	22
22 Day Ticket	44
14 Day Ticket	30
50 Journeys	50
10 Weeks	103

Data Excluded from the ODM

B.18 Some of the LENNON data has been excluded from the MOIRA2 Demand Matrix, and subsequently from the ODM.

B.19 All the products that were classified into the 'miscellaneous' ticket pot were excluded. These products were:

- Car Parking
- Railcard Sales
- Penalty/Excess Fares
- Seat Reservations
- Sleeper Supplements.

B.20 Also excluded from the analysis were all the flows that had either an Origin or Destination that did not represent a geographical location (these are mainly "I codes"), e.g.

- Rover and Ranger Tickets (except those included in the new 'Other' Infill in 2012/13);
- BritRail Tickets;
- Gate passes usually used by staff;
- Passenger Charter Discounts;
- Headquarters Input Items, other than those which can be identified as TfL or PTE.

B.21 Finally for flows that have either Origin or Destination a Private Settlement Code some are included and some are excluded.

- PTE tickets and TfL sold London Travelcard records from LENNON are removed, and replaced with an estimate of all rail travel using these tickets via 'infill's to the MOIRA2 demand matrix (refer to Chapter 3).
- PlusBus – all significant flows have been included since 2007/08, and minor flows are excluded.
- Attractions – the rail element of the significant flows have been included since 2007/08, which include:
  - Bluewater Shopping Centre
  - Alton Towers
  - Whipsnade
  - Chatsworth House

B.22 All other flows involving Private Settlement are excluded, e.g. Irish Stations.

Control Sheet

Document Title

Origin-Destination Matrix 2013/14

Document Type

Report

Client Contract/Project No.	SDG Project/Proposal No.
	22670601

Issue history

Issue No.	Date	Details
v2.0	18/12/2014	Final version for publication

Review

Originator

Jonathan Platten

Other Contributors

Reviewed by

Kevin Dadswell
Andrew Davies

Distribution

Client	Steer Davies Gleave
Office of Rail Regulation	