Estimates of Station Usage 2011/12

Methodology and Validation Report

Report

May 2013

Prepared for:
Office of Rail Regulation
One Kemble Street
London
WC2B 4AN

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

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

CONTENTS

TABLES

APPENDICES

Executive Summary

Introduction

1. This report explains the information contained within the ORR's Estimates of Station Usage data set (Station Usage 2011-12.xls) and provides guidance to the methodology followed during the process of creating this file for financial year 2011/12 and a summary of the validation checks undertaken as part of the production process.
2. The Estimates of Station Usage data set (referred to in the rest of this report as "Station Usage data set") consists of estimates of the total numbers of people:

   - Travelling from or to the station (entries & exits); and
   - Interchanging at the station (interchanges).

3. Information is given for all the national rail stations in England, Scotland, and Wales based on tickets sales data. These results are the most recent in a series produced for the ORR since 1997/98 and the spreadsheet is in a similar format to those previously provided for earlier years.
4. Station Usage data is generated from the Origin Destination Matrix (ODM), a comprehensive matrix of rail flows throughout England, Scotland and Wales, also produced by Steer Davies Gleave, and based on data produced for the MOIRA2[1] rail planning tool which itself is derived from LENNON, the rail industry's ticketing and revenue system.

Methodological Development

5. 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.
6. In the 2011/12 dataset a number of changes have been made to improve the dataset:

   - The growth rate applied to the PTE infill in the ODM has been applied at a more disaggregate level;
   - An improved PTE infill has been included for the West Midlands PTE (Centro);
   - Journeys on a number of Ranger/Rover products previously excluded has been added;
   - Entries and Exits at the central Liverpool stations have been calibrated to count data in a pilot to test the potential of use of count data in conjunction with ticket sales data.

Results

7. In total entries and exits have increased by around 6.7 % to nearly 2.47bn in 2011/12.

[1] In the 2010/11 Station Usage report MOIRA2 was referred to as MOIRA Replacement. MOIRA2 is now the adopted name for this rail planning tool.

1 Introduction

1.1 Steer Davies Gleave was appointed by the Office of Rail Regulation (ORR) to produce the Estimates of Station Usage data for 2011/12, continuing the historic series that dates back to 1997/98. This report accompanies the Estimates of Station Usage data for 2011/12 and provides details of the process and outputs used to produce the statistics on behalf of the ORR. In the rest of this report the Estimates of Station Usage data set is referred to as the "Station Usage data set."

1.2 Steer Davies Gleave are providing the ORR with an MS Excel file, "Station Usage 2011-12.xls" containing entries, exits and interchanges made at stations throughout England, Scotland and Wales, for the financial year 1st April 2011 to 31st March 2012. For the entries and exits, figures are split into the three main categories of the available ticket products (Full, Reduced, and Season).

1.3 The methodology adopted by Steer Davies Gleave in the production of the Station Usage data is consistent with that adopted by DeltaRail in the production of the Station Usage data in previous years. As part of our work we have undertaken a Methodological Review of the data and processes used to generate the Station Usage data as well as stakeholder input (see "Methodological Review Summary" report) and identified a number of areas for improvement in the data set. A number of these have been implemented in the 2011/12 data set (see Chapter 3) and others will be investigated in more detail with a view to including in future Station Usage data sets.

Use of the station usage dataset

1.4 When using the station usage data, particularly when comparing with previous years, 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

1.5 Improvements to the dataset in 2011/12 are set out in Chapter 3. 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

1.6 In the absence of a completely gated system that allows a complete recording of flows through stations or comprehensive and robust large quantities of robust count data the use of ticket sales data, LENNON, as the primary source of the station usage data set as described in the following chapter 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 E.

- **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 and will not even have point of sale information. 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;
- **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 station usage dataset as its purpose is to record bone-fide journeys on the rail network and inclusion of ticketless travel could distort business cases for new investment where these are reliant on station usage data.

1.7 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 (in particular for Station Usage statistics) the data is being pushed significantly beyond what it was originally designed for which was primarily to report and allocate revenues across train operators.

Factors which can affect reporting of entries and exits

Gating Schemes

1.8 Installation of ticket gates can significantly affect not only the usage figures at that station, but also those at neighbouring stations. The gates help to ensure that customers purchase tickets, but customers may also alter their travel patterns to avoid gated stations. We would expect travel patterns to be most affected in the months following the installation of the gates.

Change in Service Pattern

1.9 Alterations in service frequency or stopping pattern would be expected to alter station usage figures. This is particularly apparent where a group of stations along a line show similar increases or decreases. Again, this can be a long-term trend.

Ticket Issuing Facilities Changes or Product Changes

1.10 Some London stations have both underground and National Rail trains operating. LENNON does not capture tickets sold by London Underground, only those sold by TOCs. Changes in ticket facilities provided by TOCs, for example the provision of ticket machines, can therefore increase the ticket sales captured by the system.

1.11 Product changes can have an effect on passengers' purchasing patterns at rail outlets thus affecting station usage data. For example, the introduction of Oyster cards at rail outlets can affect stations inside the Travelcard boundary in the London area.

Engineering Work

1.12 Significant engineering work can alter customers' travel patterns.

Tourism

1.13 Stations near to tourist attractions may show significant changes in usage as a result of weather, promotions or other factors, which affect tourists' journeys.

New/Special Stations

1.14 Some stations serve a particular activity or business. Some fluctuation in usage of such stations is reasonable. Such activities include:

- Racecourses e.g. Newbury Racecourse
- Sports Events
- Exhibition Centre Glasgow
- Airports

1.15 In addition, where there are new stations ramp up effects can cause large demand increases over a number of years.

Trend of Growth or Decline

1.16 For stations with a history of growth or decline, it is reasonable to expect this trend to continue. There are many possible reasons for these trends, such as demographic and employment changes (new developments in the vicinity), changes in rail service levels or new stations abstracting demand.

Changes in the Sales of Individual Ticket Types

1.17 Miscoding of ticket information entered into LENNON can alter station usage results, although this would not be reflecting an actual change in customers' journeys.

Historic Events

1.18 Although not relevant for this year, there are a number of factors worth taking into account when considering generic annual data:

- Years may have been affected by industrial action such as 1994/95;
- Major incidents affecting services such as Southall, Ladbroke Grove and Hatfield;
- Major adverse weather; and
- Infrastructure changes e.g. ticket gating can significantly increases revenue - more gates have been installed in recent years which will affect the data but which does not represent higher passenger numbers.

2 Methodological Overview

MOIRA2 Demand Matrix – Base Data

Overview

2.1 All estimates of station usage, exits, entries and interchanges included in the station count dataset, are derived from the Origin Destination Matrix (ODM), also produced by Steer Davies Gleave for the ORR. The ODM itself is, in turn derived primarily from the MOIRA2 Demand Matrix.

2.2 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)[2].

Underlying Base Data - LENNON

2.3 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). However, as 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, 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.

2.4 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.

2.5 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

[2] 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.

zone, or tourist attraction. The MOIRA2 demand matrix includes the journeys and the net revenue associated with such tickets.

2.6 Data excluded from the MOIRA2 demand matrix is set out in Appendix E.

Ticket Type Definitions

2.7 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:

i) Full: all walk-up undiscounted single or return tickets, whether or not issued with a status discount (child, railcard etc);

ii) Reduced: all walk-up discounted single or return tickets, whether or not issued with a status discount (child, railcard etc);

iii) Advance: all advance-purchase tickets; and

iv) Seasons: all multi-use tickets.

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

2.8 Infills are included within the MOIRA2 demand matrix to add in the missing journeys and revenue identified in para 2.4 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.

2.9 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.

2.10 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.

2.11 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

2.12 The revenue and journeys associated with PTE Infill are estimated based on an uplift applied to the previous year's PTE Infill figures. The methodology used has changed slightly this year to use a more disaggregate growth rate and is detailed in the following chapter.

Unknown Destinations

2.13 Ticket sales do not always tell us where a passenger is travelling, for example where the Origin or Destination is a London Travelcard. As in previous years, we have converted unknown destinations into an estimate of the actual stations that passengers are travelling to. The full detail of this part of the methodology appears in Appendix D.

Interchanges Methodology

2.14 An estimate of the number of people interchanging at each station is obtained by combining the number of journeys made on each flow (from the ODM) with the information on passenger journeys taken from the Central Allocations File (CAF).

2.15 The CAF is an output of the ORCATS system which predicts passenger choices of rail route and train used, and determines the allocation of passenger revenue between TOCs. Since ORCATS is a model, the CAF contains estimates rather than actual journeys. However, it is used throughout the rail industry, so it is an appropriate source of data to use for this purpose. Since CAFs are updated with the timetable, not with financial years, no CAF will match the ticket sales data exactly. The December 2011 CAF is used in the creation of the 2011/12 Station Usage.

2.16 The CAF contains:

- Origin and destination;
- Route alternatives for each origin and destination, including all interchange points;
- Ticket type data; and
- For each flow, the proportion of passengers who choose to travel on each route alternative as calculated by the ORCATS model.

2.17 An overview of the ORCATS allocation process can be found in Appendix C.

Joint Rail & TfL Stations

2.18 Joint stations are stations which are served by both rail services and TfL services i.e. Underground or DLR. These stations can have both a TfL and a TOC ticket office, or they may have just a TfL or just a TOC ticket office. Special treatment of the ticket sales at these stations is important to ensure a realistic estimate of passengers using rail services. Passengers travelling on Underground or DLR services should not be included. These stations are identified on the station usage by the flag: London Joint Station = 'Joint TfL & TOC Station'.

2.19 Within the MOIRA2 demand matrix, an estimate of the number of travellers using rail (as opposed to other modes i.e. tube or DLR) is made. Ticket sales at the joint

stations are therefore scaled down in order to provide a better estimate of actual rail journeys.

2.20 In addition, there are a number of stations that MOIRA2 treats in the same way as Joint Stations. These are identified by the flag: London Joint Station = 'MOIRA2 Joint Station', e.g. Lewisham and Greenwich. These stations are not classified by the TOCs and TfL as a 'Joint Station', but they are the same: they are served by both rail and Underground/DLR services.

3 Methodological Changes in 2011/12

Introduction

3.1 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.

3.2 In the 2011/12 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.

Improved PTE Infill growth rate

Methodology

3.3 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.

3.4 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.

3.5 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 grown by growth rates derived from National Rail Trends data.

3.6 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 National Rail Trends data. In the construction of the 2011/12 dataset a more disaggregate set of growth rates has been applied at the PTE level based on LENNON data.

Impact of change

3.7 Growth rates applied by PTE area are shown in Table 3.2. These growth rates have enabled the capture of the specific growth characteristics of the different PTEs rather than treating them as a single group as in previous years. The average growth rate of just over 6% compares well with the 7.1% National Rail Trends journey growth rate for franchised regional operators and we have not made any adjustment to the growth rates applied to reflect the National Rail Trends number.

TABLE 3.1 PTE INFILL GROWTH RATES (2011/12)

PTE	Annual Growth	
	Journeys	Revenues
West Midlands PTE	10.05%	9.77%
Greater Manchester PTE	8.35%	13.54%
Merseyside PTE	3.23%	8.42%
Tyne and Wear PTE	-1.47%	5.70%
Strathclyde Partnership for Transport	4.62%	7.85%
South Yorkshire PTE	1.44%	8.45%
West Yorkshire PTE	5.90%	8.67%
Average	**6.08%**	**9.69%**

Source: LENNON analysis

Inclusion of revised West Midlands PTE (Centro) Infill

Methodology

3.8 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.

3.9 The infill represents a significant improvement on the current infill in the ODM, generally increasing passenger volumes across the area, 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.

Impact of change

3.10 The inclusion of the Centro infill represents 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 has resulted in an increase in entries and exits although in a small number of instances there is a decrease. A comparison of the Centro infill with the 2010/11 ODM infill is included in Table 3.2. This shows that the new infill adds approximately 5 million journeys (10 million entries and exits) compared to what would have been derived had the previous methodology been used.

TABLE 3.2 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

Methodology

3.11 In some non-PTE areas there are zonal products which are not currently 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

3.12 Journeys on these products have been 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 have been constructed using LENNON data and distributing journeys based on point of sale and the underlying reduced ticket travel distribution of the stations covered.

Impact of change

3.13 The total number of entries and exits arising from inclusion of these journeys is 760k. Table 3.3 lists the top five stations impacted most significantly:

TABLE 3.3 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

Calibration of entries and exits to count data at group stations (pilot)

Methodology

3.14 The key addition to the underlying MOIRA2 data in the construction of the station usage dataset is the breakdown of group station flows into their component stations. This is a significant task and the current methodology based primarily on sales data is becoming less robust as increasing volumes of sales are completed via the internet.

3.15 For the purposes of the 2011/12 dataset we have, therefore, conducted a pilot for stations within the Liverpool BR group of stations, using count data to allocate journeys between the stations. The stations that this impacts are:

- Liverpool Lime Street;
- Liverpool Central;
- Liverpool James Street; and
- Moorfields.

3.16 It was initially hoped to adjust journeys to these stations on a flow basis but this proved complex and only viable if applied at the most detailed flow level. Therefore, an adjustment has been applied in the station usage dataset at the station level only.

Impact of change

3.17 Count data sourced from the DfT and Merseytravel enabled us to calculate the split of demand between the central Liverpool stations as shown in Table 3.4. We then used these percentages to divide total central Liverpool demand, as calculated by the station usage process, between the central Liverpool stations. The same splits have been applied across all ticket types.

TABLE 3.4 MODIFICATION OF CENTRAL LIVERPOOL STATION USAGE DATA

Station	2011/12 Entries and Exits old methodology	Implied split between stations	Implied split between stations from counts	Adjusted Liverpool station entries and exits
Liverpool Lime Street	11,882,144	32%	37%	13,835,314
Liverpool Central	17,497,878	47%	38%	14,209,241
Liverpool James Street	3,524,654	9%	8%	2,991,419
Moorfields	4,488,064	12%	17%	6,356,766

Other methodological variations

3.18 In the 2010/11 station usage calculations the generic methodology for separating out group stations was not followed for the following BR groups due the necessary data not being available in LENNON:

- Wigan BR
- Warrington BR
- Manchester BR

3.19 It has been possible to use the generic methodology for Manchester and Warrington in 2011/12 but still not for Wigan where we have maintained the same split of journeys between Wigan North Western and Wigan Wallgate as in 2010/11 at a flow and route code level.

4 Summary of Results

4.1 The following table gives the total number of entries, exits, and interchanges made over the whole network for 2010/11, compared with the previous year.

TABLE 4.1 ENTRIES, EXITS AND INTERCHANGES FOR 2010/11 – 2011/12

Year	Entries	Exits	Entries & Exits	Interchanges
2010/11	1,156,896,520	1,156,896,522	2,313,513,546	165,529,865
2011/12	1,234,360,071	1,234,360,071	2,468,720,142	202,130,946

4.2 Overall, the increase in entries and exits is around 6.7% in 2011/12, compared with the previous year.

Overview of the Entries and Exits Results

4.3 In this section we set out a summary of the overall entries and exits results. The spreadsheet contains entries and exits results for 2,533 stations, compared with 2,531 last year. The tables below show the stations no longer in the Station Usage data this year, and the new stations that have been opened.

4.4 In 2011/12, two stations were added and one station, Smitham, was renamed as Coulsdon Town and thus appears as both an addition and a removal.

TABLE 4.2 STATIONS IN 2010/11 BUT NOT IN 2011/12

NLC	Name	Note
5382	Smitham	Renamed “Coulsdon Town”

TABLE 4.3 STATIONS IN 2011/12 BUT NOT IN 2010/11

NLC	Name	Note
5382	Coulsdon Town	Previously named “Smitham”
4787	Southend Airport	New station
7501	Buckshaw Parkway	New station

4.5 The table below shows data for the ten stations with the highest numbers of entries and exits for 2011/12.

TABLE 4.4 TOP 10 STATIONS BASED ON 2011/12 ENTRIES AND EXITS

Rank This Year	NLC	Station Name	2011/12 Entries & Exits	2010/11 Entries & Exits	Change	Rank Last Year
1	5598	Waterloo	94,045,510	91,750,382	3%	1
2	5426	Victoria	76,231,290	73,573,492	4%	2
3	6965	Liverpool Street	57,106,502	55,769,423	2%	3
4	5148	London Bridge	52,634,024	51,478,131	2%	4
5	5143	Charing Cross	38,004,790	37,222,298	2%	5
6	1444	Euston	36,608,546	34,073,413	7%	6
7	3087	Paddington	33,736,546	32,200,316	5%	7
8	1127	Birmingham New Street	31,213,842	24,686,632	26%	10
9	6121	King's Cross	27,874,732	26,254,644	6%	8
10	9813	Glasgow Central	26,639,418	24,950,987	7%	9

4.6 The total journeys made at one of the top ten stations account for a total of 474 million, 4.9% more than the 452m journeys made at the top ten stations of last year although it should be noted that this includes a substantial increase for Birmingham New Street due to the inclusion of the improved Centro area PTE infill. If Birmingham New Street is excluded the increase is 3.6%. The top ten stations account for 19% of all entries and exits, marginally less than the 20% share in 2010/11.

Merseyside

4.7 Merseyside as a whole and some stations within or on the edge[3] of the Merseyside area are showing unexpected decreases in entries and exits (Merseyside is showing a 1% decrease compared with a circa 2% increase last year). Investigation indicates that this is related to a switch between travel on PTE and non-PTE tickets in some areas which is not captured in the current PTE infill growth methodology. Users should bear this in mind when using the data. This is an area which is a focus for improvement in the next annual update.

Overview of the Interchanges Results

4.8 In all, around 202 million interchanges are estimated to have been made among National Rail operated services (interchanges between rail and tube or other modes are excluded except for cross-London journeys). This is an increase of 22% compared to the 2010/11 results (165.5 million). The ten top stations are listed in the table below.

[3] Three of the most affected stations are Ormskirk, Aughton Park and Town Green

4.9 Approximately half of the increase in interchanges is driven by an improvement in the treatment of journeys involving non-national rail portions (e.g. cross-London trips involving the Underground). This has a particular impact on interchanges at London Termini.

TABLE 4.5 TOP 10 STATIONS BASED ON THE INTERCHANGES MADE FOR 2011/12

Ranking 2011/12	NLC	Station Name	2011/12 Interchanges	2010/11 Interchanges	Change	Ranking 2010/11
1	5595	Clapham Junction	21,609,997	20,667,636	5%	1
2	5598	Waterloo	9,488,822	5,772,501	64%	4
3	5426	Victoria	9,156,710	4,800,979	91%	5
4	5148	London Bridge	8,741,879	7,346,732	19%	2
5	5355	East Croydon	6,340,937	7,113,300	-11%	3
6	1127	Birmingham New Street	5,117,520	4,319,983	18%	6
7	1444	Euston	3,831,564	1,808,443	112%	10
8	2968	Manchester Piccadilly	3,795,951	2,624,292	45%	8
9	3149	Reading	3,793,740	2,898,671	31%	7
10	1555	St.Pancras	3,675,849	2,158,903	70%	9

4.10 Interchanges occurred at 538 stations in 2011/12 compared to the 536 stations in 2010/11. Stations appearing for the first time in 2011/12 and those not seen this time are listed below.

TABLE 4.6 CHANGES IN INTERCHANGE STATIONS IN 2011/12

	Interchanges		Reason
	2011/12	2010/11	
New			
Cannonbury	72,868	0	Extension of London Overground services
Shoreditch High Street	46,345	0	
Pontefract Tanshelf	160	0	
Old			
Boston	0	3,702	
Dalston (Kingsland)	0	187,533	

4.11 The numbers in this table are estimated numbers for actual passenger interchanges made during the year.

4.12 it is important to note that interchanges can change significantly from year to year for a variety of reasons. Factors such as new service patterns and changes in journey times play a part. The number of interchanges is based on the rail industry ORCATS model, which predicts passenger choices of rail route and trains used. Refer to Appendix C for more information on the ORCATS allocation process.

5 Validation

Introduction

5.1 The production of the 2011/12 dataset has seen the revision of the checking and validation processes associated with the Station Usage data. Checks undertaken on the station usage dataset encompass a number of elements, including:

- Investigation of large increases and decreases for individual stations
- Checks at different geographical levels
- Validation against alternative data sources

Data Checks

Large increases and decreases

5.2 Table 5.1 shows the 10 stations with the largest increases in total flow for stations with more than 10,000 entries and exits.

TABLE 5.1 TOP 10 INCREASES

NLC	Station Name	2010/11 Entries and Exits	2011/12 Entries and Exits	Increase (%)	Reason
990	Armadale	11,170	126,086	1029%	Station opened in March 2011 - first full year of operation
5036	Doleham	3,894	38,666	893%	Large increase consistent with previous year and appears to be driven by timetable improvements.
992	Caldercruix	11,120	90,976	718%	Station not opened until February 2011 - first full year of operation
3498	Lelant Saltings	17,224	101,284	488%	Improved data due to inclusion of St Ives branch Ranger tickets (additional 79,000 entries and exits)
9757	Drumgelloch	58,550	269,172	360%	Station re-opened in March 2011
3542	Carbis Bay	55,334	206,736	274%	Improved data due to inclusion of St Ives branch Ranger tickets (additional 142,000 entries and exits)
991	Blackridge	12,394	43,258	249%	Station only opened in December 2010 - demand will still be ramping up
3092	Kensington Olympia	2,311,792	5,936,984	157%	London Overground demand ramp-up and increased demand following reduction of District line services
1441	Canonbury	772,976	2,102,340	172%	Improved North London Line and East London Line services causing large passenger increase
9668	Bogston	25,744	62,992	145%	Cause of demand increase not clear

5.3 Table 5.2 shows the 10 stations with the largest decreases in total flow for stations with more than 10,000 entries and exits.

TABLE 5.2 TOP 10 DECREASES

NLC	Station Name	2010/11 Entries and Exits	2011/12 Entries and Exits	Decrease (%)	Reason
9982	Helensburgh Upper	23,466	14,198	-39%	Partially driven by drop in journeys to Glasgow Queen Street which form the main component of journeys from this station
6268	Wainfleet	64,676	41,578	-36%	Reason for change unclear - driven by large decrease in season ticket demand
4617	Cosford	85,072	61,378	-28%	Station rebuilt during year
5237	Maidstone Barracks	160,029	117,538	-26%	Continuing effect of HS1 services moving demand to Maidstone West from other Maidstone stations
2283	Town Green	193,200	149,306	-23%	Linked to Merseyside issues discussed in para 4.7
9544	Prestwick Internat'nl Airport	432,334	336,982	-22%	Reflects lower passengers at the airport
2242	Liverpool Central	17,958,028	14,209,241	-21%	Reflects central Liverpool station adjustment
9555	Stranraer	57,276	45,388	-21%	Decrease due to cessation of ferry services at Stranraer
4643	Stourbridge Town	592,838	471,220	-21%	Decrease due to improved demand estimation in Centro area (108,000 less journeys in Centro infill compared to 2010/11 Centro infill data).
5762	Lympstone Commando	47,660	38,616	-19%	Continues recent decline in passenger numbers

5.4 Within the 2011/12 dataset two new flags have been introduced identifying:

- Stations with more than 10,000 entries and exits a year where entries and exits have increased or decreased by more than 10%
- Stations with less than 10,000 entries and exits a year where entries and exits have increased or decreased by more than 25%

5.5 These flags have been used to identify stations where further investigation should be carried out to ensure, where possible, the reported changes reflect reality. The limits set are demanding (10% of 10,0000, for example could represent just two extra season ticket holders per year)and investigations have been focussed on the most significant changes but where obvious explanations for less significant changes are available these have been included in the Station Usage dataset.

5.6 In total 668 stations were captured by one of the two flags. Whilst a large number, it is less than would have been captured by similar checks on the previous year dataset (comparable number is 840).

5.7 Whilst reasons for large changes at some stations are specific to that station, in many instances there are groups of stations where there is a common cause for the changes seen. In Table 5.3 we have grouped reasons for large changes into a number of clusters, with the number of stations in each category. The table shows that the highest identifiable drivers of significant change are the continuing increase in patronage on London Overground routes e.g. continued ramp-up of demand following opening of East London Lone and the improved Centro area PTE infill. In total, the types of reasons summarised in Table 5.3 account for 61% of all flagged stations when we include stations where there is a trend of high growth, albeit not for obvious reasons.

TABLE 5.3 SUMMARY OF LARGE CHANGES

Reason	Note	Stations affected
Centro Infill	Methodological improvement	64
St Ives Line		4
London Overground effect		70
Timetable Improvement		21
Airdrie - Bathgate		14
New station demand ramp-up		10
HS1		8
Merseyrail infill issues		3
Engineering works		3
High growth trend		209

Checks at different geographical levels

5.8 It is possible that in certain areas changes at the individual station level might not be large enough to be flagged but as a group the results might be unexpected. For this reason we have carried out some checks at a number of levels of detail. In this section we summarise the station count data for the following aggregations of data:

- PTE area;
- Government Office Region (GOR); and
- Station Facility Owner (SFO).

TABLE 5.4 ENTRIES AND EXITS BY PTE AND LONDON TRAVELCARD AREA

PTE	Entries and Exits		Growth
	2010/11	2011/12	

London Travelcard Area	1,081,196,831	1,171,658,378	8.37%
Greater Manchester	63,665,262	68,717,462	7.94%
Merseyside	92,414,145	91,487,191	-1.00%
South Yorkshire	18,941,094	19,739,157	4.21%
Strathclyde	110,398,179	117,417,817	6.36%
Tyne & Wear	8,731,112	8,954,523	2.56%
West Midlands	105,651,484	121,834,750	15.32%
West Yorkshire	60,886,855	62,443,056	2.56%

5.9 This table shows sensible increases by PTE with the high growth rate seen in the West Midlands a direct result of the improved Centro infill. The decrease for Merseyside is discussed earlier in para 4.7.

TABLE 5.5 ENTRIES AND EXITS BY GOVERNMENT OFFICE REGION

GOR	Entries and Exits		Growth
	2010/11	2011/12	
London	1,069,338,251	1,159,311,691	8.41%
South East	341,527,637	355,569,907	4.11%
East	174,132,068	181,725,667	4.36%
South West	63,287,438	67,962,983	7.39%
East Midlands	37,247,340	38,461,207	3.26%
West Midlands	104,146,206	120,177,980	15.39%
North East	19,501,310	20,001,874	2.57%
North West	194,483,104	199,946,647	2.81%
Yorkshire And The Humber	99,565,171	102,783,800	3.23%
Wales - Cymru	45,944,370	47,134,704	2.59%
Scotland	164,340,651	175,854,054	7.01%

5.10 Changes by GOR are, again, all within the range of what would be expected with highest growth in London (driven partly by London Overground) and the South West.

TABLE 5.6 ENTRIES AND EXITS BY STATION FACILITY OWNER

SFO	Entries and Exits		Growth
	2010/11	2011/12	
Arriva Trains Wales	53,647,288	54,906,734	4.3%
c2c	44,534,718	47,890,560	13.8%
Chiltern Railways	33,586,328	37,267,690	8.1%
East Coast	32,308,884	33,643,264	5.0%
East Midlands Trains	38,374,508	39,472,600	3.5%
First Capital Connect	112,441,495	115,994,692	5.5%
First Great Western	112,344,542	120,318,740	10.1%
First ScotRail	118,666,068	125,888,884	4.1%
First TransPennine Express	21,470,589	22,308,272	9.7%
Glasgow Prestwick Airport	432,334	336,982	-18.8%
London Midland Trains	68,070,475	76,727,936	7.6%
London Overground	69,608,231	99,806,578	66.9%
London Underground	38,375,583	50,111,026	86.9%
Merseyrail	75,519,049	72,112,632	0.6%
National Express East Anglia	149,781,128	161,490,108	10.3%
Network Rail	601,917,586	629,047,816	6.6%
Northern Rail	106,573,058	116,328,320	8.9%
South West Trains	253,096,633	266,306,990	6.5%
South West Trains (Island Line)	1,627,348	1,667,834	6.6%
Southeastern	175,640,492	180,237,504	5.3%
Southern	168,114,537	178,310,310	6.5%
Virgin Trains (West Coast)	37,382,672	38,528,568	7.2%

5.11 Changes at the SFO level are also within reasonable bounds. The large decrease for Glasgow Prestwick Airport reflects a corresponding drop in passengers using the airport. The large increase for London Overground is consistent with growth on the Overground network.

Validation against alternative data sources

Comparison with ORR journey data on the ORR data portal

5.12 The ORR produces journey data by sector and TOC and makes this available on the ORR website via its data portal through a separate data analysis exercise[4]. Growth from 2010/11 to 2011/12 from this data was 7.8% at the national level for franchised TOCs. The station usage data shows an increase of 6.7% over the same period, within 1.5% of the data portal data.

Comparison with PIXC data

5.13 The DfT collects count data for major cities throughout the UK. The method of collection means that for through stations it is often not possible to calculate boarders and alighters but for terminal stations this is usually possible. Using data provided by the DfT we have compared growth rates at the major London termini covered by the count data with those seen in the calculated station usage data. The only regional station where the comparison is possible using DfT data is Manchester Victoria.

TABLE 5.7 COMPARISION OF STATION USAGE AND PIXC GROWTH RATES 2010/11 – 2011/12

Station	Station usage growth rate (all day)	PIXC growth rate (AM + PM Peak)
Euston	7%	8%
Fenchurch Street	2%	6%
King's Cross	6%	6%
Liverpool Street	2%	7%
Marylebone	9%	10%
Moorgate	9%	11%
Paddington	5%	-5%
Victoria	4%	2%
Waterloo	3%	2%
Manchester Victoria	56%	8%[1]

[1] Note the Manchester Victoria counts numbers are start of service to 21:00.

5.14 Conclusions that can be drawn from the comparison for the London stations are limited because the count data is peak only. However, the growth rates are reasonably comparable except for Paddington and Liverpool Street and Fenchurch Street. In most instances the station usage data growth rate is lower than the count growth rate which could well reflect a higher growth rate in peak compared to off-peak demand. Finally, the increase in the station usage dataset for Paddington appears more plausible than the count decrease.

[4] Formerly this formed part of the National Rail Trends publication

5.15 The growth rate obtained for Manchester Victoria is much large than that seen in count data. However, there are a number of reasons for believing that the current station usage number is an improvement on the previous year:

i) Issue with previous year means that we are effectively capturing two years change;

ii) We are aware of increased revenue protection activity which is likely to have resulted in increased ticket sales at Manchester Victoria and on commuter flows into Manchester Victoria; and

iii) A comparison with count data suggests the Manchester Victoria station usage number is now closer to the absolute count data than previously.

APPENDIX

A

HISTORICAL METHODOLGICAL CHANGES

A1 HISTORICAL METHODOLOGICAL CHANGES

A1.1 In the five years prior to the latest dataset a number of improvements were made to the ODM and Station Usage methodology which are described in this section. In addition since 2009/10 Oyster PAYG data has been included in the ODM which represented a significant improvement to the estimates for rail travel across London.

Historical methodology changes

A1.2 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 - t*his 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.

A1.3 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.

A1.4 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

A1.5 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.

A1.6 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.

APPENDIX

B

STATION USAGE FILE DEFINITION

B1 STATION USAGE FILE DEFINITION

B1.1 The Station Usage spreadsheet (Station Usage 2011-12.xls) lists the entries, exits and interchanges made at stations throughout England, Scotland and Wales in the financial year 2011/12 (1st April 2011 to 31st March 2012). It also gives details about the entries and exits for different ticket categories. It contains data on entries and exits made at rail stations by passengers using the rail network. The fields included in the Station Usage data set are:

APPENDIX TABLE B.1 STATION USAGE FILE

Field	Description
Station (Name, NLC, TLC)	Station Name, NLC: National Location Code, TLC: Three Letter Code
District, Country, Region, NUTS2 Code and NUTS2 Spatial Unit for the Station	District, County, Region, NUTS2 Station's geographical location
Station Facility Owner (SFO)	The company that is the station facility owner (provided by Network Rail in 2008)
Station Group	Name of the Group where applicable. The user of this data may wish to filter on the 'Station Group' column, or create pivot tables, to investigate the results at a group level
PTE Urban Area Station	Stations within the urban areas covered by PTE services are identified with a flag: 'PTE Urban Area Station'
London Travelcard Area	Stations with the urban areas covered by PTE services and TfL services are identified with a flag: 'London Travelcard Area Station'
London Joint Stations	Joint stations which are served by both rail services and TfL services are identified with a flag: 'Joint TfL & TOC Station'
Entries (Full, Reduced, Season, Total)	Entries made at the stations split by ticket categories and in total
Exits (Full, Reduced, Season, Total)	Exits made at the stations split by ticket categories and in total
11/12 Entries & Exits	Sum of Entries and Exits for 2011/12
10/11 Entries & Exits	Sum of Entries and Exits for 2010/11
11/12 Interchanges	Total Interchanges made for 2011/12
Large station Flag	Flags change in Entries and Exits greater than 10% for stations with over 10,000 Entries and Exits
Small station Flag	Flags change in Entries and Exits greater than 25% for stations with under 10,000 Entries and Exits
Explanation of large change	Identified reason(s) for large changes for flagged stations
Sources	Links to source(s) of information where appropriate

Regions, Counties and Districts

B1.2 For all rail stations, the District, County, Region and NUTS2 Region & Code are provided for the origin and destination to describe the geographical location.

B1.3 The source of this data is:

- District or the Unitary Authority – ATOC (dated January 2008) and ORR (dated January 2008)
- District, County & Region – ONS[5] website (dated January 2008)
- NUTS2 Code and Description – ORR (dated January 2010)

[5] http://www.statistics.gov.uk/geography/geographic_area_listings/administrative.asp#04

APPENDIX

C

OVERVIEW OF THE ORCATS ALLOCATION PROCESS

C1 OVERVIEW OF THE ORCATS ALLOCATION PROCESS

C1.1 This section gives an outline of the Central Allocations File (CAF), which is used in producing the interchange figures, and the ORCATS process which is used to create the CAF.

C1.2 Most of the train tickets that are sold are inter-available – the customer has a choice of routes and operators. For example, when a customer buys a ticket to travel from Leicester to Leeds, that customer may travel on various combinations of East Midlands Trains, East Coast, CrossCountry Trains and Northern, and may interchange at Doncaster, Sheffield, Derby or Nottingham. LENNON captures the sale of the ticket, but unless the ticket has stringent route restrictions, the route actually taken by the customer is not recorded.

C1.3 The route taken by any particular customer may never be known, but some route options are more attractive than others. The customer is more likely to choose a faster, more frequent service than a slower, less frequent one. This likelihood can be translated into the proportions of customers choosing each route option, on a particular flow. (A 'flow' represents all journeys from a given origin station to a given destination station, irrespective of the route taken.) The revenue received from all customers on that flow should be split between different operators to reflect the proportion of customers which each operator carried.

C1.4 ORCATS was developed to model the choice made by the customers, and to allow revenue to be split between operators. It applies passenger choice modelling to the train timetable, to determine the relative attractiveness of different route alternatives. It then weights the results by journey mileage.

C1.5 For any given timetable, ORCATS works out the possible routes between each origin and destination, and calculates the percentage of the passengers that are expected to choose each route based on the services in that timetable.

C1.6 The output from ORCATS is the Central Allocations File (CAF). This lists the proportion of journeys on each flow (or origin-destination pair) estimated to be made by each route alternative. For journeys involving interchanges, each leg of the journey is listed. By combining this information with the ODM data, which contains journeys for all flows, the number of interchanges occurring at individual stations has been estimated.

APPENDIX

D

METHODOLOGY: NON-STATION TICKETS

D1 METHODOLOGY: NON-STATION TICKETS

D1.1 Ticket sales do not always tell us where a passenger is travelling. Ticket sales can be divided into the seven categories listed in table below. Ticket sales data has been converted into an estimate of the actual stations that passengers are travelling from/to.

D1.2 The processing of ticket sales data is undertaken in the creation of the MOIRA2 demand matrix, and then subsequently in the creation of the ODM. For each of the flow categories, the table below states where the flow is processed: MOIRA2 or ODM.

APPENDIX TABLE D.1 CATEGORISATION OF TICKET SALES IN LENNON

Flow Category	Description	Processing
Category 1	Origin and Destination Stations Known	No processing required
Category 2	Origin or Destination a Group Station (excl. London BR)	ODM
Category 3	Origin or Destination is London Terminals	ODM
Category 4	Origin or Destination a London Travelcard including Zone 1	ODM
Category 5	Origin or Destination a London Travelcard excluding Zone 1	MOIRA2 Demand Matrix
Category 6	Origin or Destination a London Travelcard Boundary Zone	MOIRA 2 Demand Matrix
Category 7	Non-National Rail Stations	MOIRA 2 Demand Matrix

Category 1 – Origin and Destination Stations Known

D1.3 Both the origin and destination were known stations so no further processing is required for such flows.

Category 2a – Origin or Destination a Group with all Stations Having a Ticket Office

D1.4 In 2005/06 all origins or destinations that were a group station (with the exception of London BR) were changed to the major station within the group. For example, all ticket sales to or from Reading BR were recoded to Reading. This was clearly over-simplistic.

D1.5 In 2006/07 the ODM was based on the journeys from ticket sales to the individual stations within a group. We assumed that passengers travelling **to** the stations in a group would act in the same way as passengers travelling **from** the stations in that group. We believed that this was, in general, a valid assumption to make, and no bias would be introduced into the journey figures.

D1.6 From 2007/08 onwards this process is still used where all stations in the group have ticket offices, so that the relative flows from the individual stations are credible.

D1.7 For example, in 2006/07 the journeys between stations in the 'Manchester BR' group and Crewe and vice-versa are shown by the column "jnys" in the table below. First the proportion of journeys **from** each of the individual Manchester stations **to** Crewe is determined, as shown in column "%split".

D1.8 Then these proportions are applied to both the 'Manchester BR to Crewe' and 'Crewe to Manchester BR' flows, giving the breakdowns to individual stations shown in column 'BR portion'. These are added to the base values to give "Total Journeys", before the 'Manchester BR to Crewe' and 'Crewe to Manchester BR' flows are deleted, to avoid double counting. The slight discrepancy between the Grand Totals is due to rounding error.

APPENDIX TABLE D.2 EXAMPLE OF BREAKING DOWN JOURNEYS TO/FROM A BR GROUP OF STATIONS

Orig	Dest	Origin Name	Destination Name	Jnys	%Split	BR portion	Total Jnys
2963	1243	DEANSGATE	CREWE	83	**0.32%**	85	168
2966	1243	MANCH OXF RD	CREWE	5,464	**21.03%**	5580	11,044
2968	1243	MANCH PICC	CREWE	19,733	**75.95%**	20152	39,885
2970	1243	MANCH VICT	CREWE	700	**2.69%**	714	1,414
0438	1243	MANCH BR	CREWE	26,533		**Remove**	
1243	2963	CREWE	DEANSGATE	207		1478	1,685
1243	2966	CREWE	MANCH OXF RD	2,262		97287	99,549
1243	2968	CREWE	MANCH PICC	8,017		351349	359,366
1243	2970	CREWE	MANCH VICT	343		12464	12,807
1243	0438	CREWE	MANCH BR	462,578		**Remove**	
		Grand Total:	525,920			525,918	

D1.9 The above methodology has been applied to all flows with more than 1,000 journeys in total, based on sales data, leaving the individual group stations (i.e.

not including the 'BR Group NLC to destination' flow). For the smaller flows an average split is applied based on the flow with more than 1,000 journeys.

Category 2b – Origin or Destination a Group with some Stations Having no Ticket Office

D1.10 For this class of stations the above process breaks down because the proportion of journeys **to** the group stations with no ticket offices will tend to be estimated as zero because the sales **from** those stations are necessarily zero. For these groups bespoke methodology has tended to be used based on the best available data. This year the same splits between the affected group stations have been maintained at a origin and destination and route code level as was estimated in the 2010/11 dataset.

Category 3 – Origin or Destination is London BR

D1.11 This category contained all flows that had London BR as either the origin or destination. In order to assign an appropriate London station on flows where either the origin or destination is London BR (NLC=1072) or a London Travelcard involving Zone 1, we analysed responses from the 2001 London Area Travel Survey (LATS). For journeys from any given station, we established the percentage of passengers using each London terminus.

D1.12 For example, if the flow was from Ashford International to London BR, we used our pre-generated table showing the percentage spilt between the alternative London termini for passengers starting at Ashford International. From this we apportioned the exits between London Bridge, Charing Cross, Victoria and other London termini.

D1.13 Stations with small sample sizes were removed from the 2001 LATS data. Where there was insufficient data in the 2001 LATS to generate the split for a particular station, a similar process with the Non London Groups methodology was applied. Firstly for all the flows with more than 1000 journeys leaving London BR and having as a destination the particular station we used split factors as above. However, if the sum of journeys was less than 1000 we assigned to the flow the top origin from the London BR stations.

Category 4 – Origin or Destination a London Travelcard including Zone 1

D1.14 All origins and destinations that were London Travelcard Zones that include Zone 1 were converted to 'London BR' under the assumption that they will travel to the same stations as point-to-point passengers and then transfer to another mode. The methodology set out above for Category 3 was then applied.

Category 5 – Origin or Destination a London Travelcard excluding Zone 1

D1.15 This category contained all Travelcards that did not include Zone 1, for example Zone R2345 London.

D1.16 For flows with origin or destination a London Travelcard (excluding zone 1) we use a set of assumptions based on survey responses from the 2001 LATS. They use the starting station to work out which stations it is possible for the passenger to be travelling to, and also give the proportion of passengers travelling to each of these stations. This is based on the assumption that a passenger holding a Zones 2-6 Travelcard would travel as far as Zone 2.

D1.17 This processing is undertaken during the production of the MOIRA2 demand matrix.

Category 6 – Origin or Destination a Boundary Zone

D1.18 All origins and destinations that were a London Travelcard Boundary Zone were converted to 'London Travelcard including Zone 1' under the assumption that a passenger travelling from or to a Boundary Zone will hold a Travelcard that includes Zone 1. The methodology set out above for Category 3 was then applied.

D1.19 This processing is undertaken during the production of the MOIRA2 demand matrix.

Category 7 – Non-National Rail Stations

D1.20 This final category contains all those flows in the original ticket sales data that do not fall into one of the above categories. Refer to Appendix E for a detailed description of this data and what has been included and excluded from the ODM.

D1.21 This processing is undertaken during the production of the MOIRA2 demand matrix.

APPENDIX

E

STATION USAGE DATASET LIMITATIONS

E1 STATION USAGE DATASET LIMITATIONS

Limitations of the LENNON data

E1.1 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.

E1.2 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.

E1.3 We have used similar information extensively in the last ten years or more, and have found the data to be reliable, particularly when examining the data at an aggregated level.

E1.4 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

E1.5 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.

E1.6 Human error at the point the ticket sale is entered into the input machines will also produce invalid data in LENNON.

Travelcards

E1.7 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.

E1.8 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

E1.9 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

E1.10 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

E1.11 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.

E1.12 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.

E1.13 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

E1.14 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

E1.15 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 Table 9-1 overleaf.

E1.16 Ticket periods of other lengths are converted to a number of journeys using a proportion of the monthly journey factor.

E1.17 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.

E1.18 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 E.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 Station Usage

E1.19 Some of the LENNON data has been excluded from the MOIRA2 Demand Matrix, and subsequently from the ODM.

E1.20 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.

E1.21 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 2011/12)
- 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

E1.22 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 2).
- 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

E1.23 All other flows involving Private Settlement are excluded, e.g. Irish Stations.

CONTROL SHEET

Project/Proposal Name	Estimates of Station Usage 2011/12
Document Title	Methodology and Validation Report
Client Contract/Project No.	Click here to enter text.
SDG Project/Proposal No.	22513001

ISSUE HISTORY

Issue No.	Date	Details
V2.0	02/05/2013	Final – Table 5.6 updated

REVIEW

Originator	Kevin Dadswell	
Other Contributors	Andrew Davies; Andy Dundas	
Review by:	Print	Andrew Davies
	Sign	

DISTRIBUTION

Client:	Office of Rail Regulation
Steer Davies Gleave:	