Final Report
April 2019

Journeys per Season Ticket Study

Department for Transport
Our ref: 23257401
Client ref: PPRO 04 101 130

steer wsp

Final Report
April 2019

Journeys per Season Ticket Study

Prepared by:

Steer
28-32 Upper Ground
London SE1 9PD

WSP
70 Chancery Lane
London WC2A 1AF

+44 20 7910 5000
www.steergroup.com

Prepared for:

Department for Transport
Great Minster House
33 Horseferry Road
Westminster
London
SW1P 4DR

Client ref: PPRO 04 101 130
Our ref: 23257401

Steer has prepared this material for Department for Transport. This material may only be used within the context and scope for which Steer has prepared it and may not be relied upon in part or whole by any third party or be used for any other purpose. Any person choosing to use any part of this material without the express and written permission of Steer shall be deemed to confirm their agreement to indemnify Steer for all loss or damage resulting therefrom. Steer has prepared this material using professional practices and procedures using information available to it at the time and as such any new information could alter the validity of the results and conclusions made.

Contents

Executive Summary ..... **i**

1 Introduction **1**

Context of the Research 1

Structure of Report 4

2 Methodological Approach **5**

Methodology overview 5

Data Sources 6

Summary of Assumptions 11

3 Stage 1 – Station Selection **12**

Selection Criteria 12

Station Selection 13

4 Stage 2 – Flow Selection **18**

Selection Process and Criteria 18

Flow Categorisation 19

Summary of Flows 20

5 Stage 3 – Calculation of Journeys per Ticket estimates **25**

Approach 25

Methodology 26

Data Miner Tool 26

Weekly Journeys per Ticket 27

Annual Journeys per Ticket 30

Monthly Journeys per Ticket 33

6 Stage 4 – Validation of Gatelines and Flows **36**

Criteria for Validation 36

High-Level Validation 37

7 Results and Key Findings **38**

Key Findings 38

Confidence Intervals 42

8 Quality Assurance 44

Overview 44

Calculation of Journeys per Ticket estimates 44

Spreadsheet Modelling Best Practice (SMBP) 45

Flow Selection, Categorisation and Validation 46

Internal Meetings and Reviews 46

Quality Assurance Actions 46

9 Limitations and Caveats to our Analysis 48

Risks and Mitigations 48

10 Conclusions 53

Recommendations 54

Appendices 1

Figures

Figure 1-1: Proportion of Commuting Journeys in a Week – Full and Part Time Workers, across All Transport Modes 2

Figure 1-2: Growth in passenger journeys on Season tickets and All ticket types 3

Figure 2-1: Stages of our Methodology 5

Figure 3-1: Gated stations in Great Britain and NLC zones including selected stations for analysis 16

Figure 4-1: Map of all stations and flows used in the analysis 21

Figure 4-2: Proportion of flows by category 23

Figure 5-1: Flowchart of overall approach for calculating Journeys per Ticket from gateline and LENNON data 25

Figure 5-2: Weekly Season Tickets Observations 27

Figure 5-3: Weekly Season Ticket Journey per Ticket Calculation 29

Figure 5-4: Monthly and Annual Ticket Observations 30

Figure 5-5: Monthly LENNON Data Matrix 31

Figure 5-6: Mask on Observations to leave exclusively Annual Ticket observations 31

Figure 5-7: Summary of approach to calculating Annual Journeys per Ticket 32

Figure 5-8: Normalised Annual 'Busyness' 34

Figure 5-9: Annual busyness shape and Monthly Ticket observations 35

Figure 7-1: Map of Weekly Journeys per Ticket by line of route from London Marylebone Station 40

Figure 7-2: Correlation between Weekly Journeys per Ticket and Fare Ratio (Anytime Return and Weekly Season) 41

Figure 7-3: Correlation between Weekly Journeys per Ticket and distance of journey 41

Figure 7-4: Correlation between Monthly Journeys per Ticket and distance of journey 42

Figure 7-5: Correlation between Annual Journeys per Ticket and distance of journey 42

Figure 8-1: Weekly JpT for Bi-Directional Flows 45

Figure 10-1: Passenger journeys growth, with and without inferred Journeys per Ticket 54

Tables

Table 2-1: Data sources used in analysis 7

Table 2-2: Strengths and Weaknesses of Gateline Data 7

Table 2-3: LENNON Query 1 8

Table 2-4: LENNON Query 2 8

Table 2-5: Reported Fields in LENNON Queries 1 and 2 9

Table 2-6: LENNON Query 3 9

Table 2-7: LENNON Query 4 9

Table 2-8: Reported Fields in LENNON Queries 3 and 4 9

Table 2-9: Summary of our assumptions in the analysis of Journeys per Ticket estimates 11

Table 3-1: NLC Zones 13

Table 3-2: Final Station Selected 15

Table 3-3: Season ticket entries and exits by station (2016/17 ORR Estimates of Station Usage) 17

Table 4-1: LENNON filters and operation, as specified on flow selection 18

Table 4-2: LENNON fields specified on flow selection 18

Table 4-3: Mapping of Flows to Flow Types 20

Table 4.4: Flows selected by distance and flow type category, prior to validation 22

Table 4-5: RUDD Proportions by Category 24

Table 6-1: Description of flow exclusion criteria 36

Table 6-2: Summary of number of flows excluded by reason 37

Table 6-3: Comparison of selected flows compared to RUDD Season ticket totals for each PDFH flow type category 37

Table 6-4: Comparison of selected flows compared to RUDD totals for each journey distance band 37
Table 7-1: High level Journeys per Ticket figures 38
Table 7-2: Weekly Journeys per Ticket results 39
Table 7-3: Monthly Journeys per Ticket results 39
Table 7-4: Annual Journeys per Ticket results 39
Table 7.5: Confidence Intervals for JpT estimates, at a 95% level of confidence 43
Table 8.1: Summary of Quality Assurance Actions 47
Table 9.1: Summary of main risks and mitigations in our analysis 48

Executive Summary

The aim of this study has been to estimate Journeys per Ticket ("JpT") figures that reflect a recent view of the frequency of travel on rail Season tickets. The figures used in LENNON, linking number of journeys to Season ticket issues, were derived over 30 years ago using evidence which is unlikely to reflect current commuting behaviour.

There is evidence from multiple sources showing that commuters now travel less frequently and, when they do so, they increasingly favour walk-up tickets (Anytime or Off-Peak) over Season tickets.

In addition, we have looked to derive disaggregated JpT figures depending on the key factors that affect them, such as distance, flow type and ratio between Anytime and Season fare price.

We have used gateline data, provided by Cubic, to estimate these new JpT figures based on actual passenger behaviour. Gateline data provides a wealth of information which includes ticket type and the day of use, and is available for a large time period and a significant number of stations. This data source has some limitations, but the methodology (see Chapters 2 to 6 for further detail on our methodology) we have used has been developed to mitigate these:

- Gateline data does not in general identify tickets uniquely, meaning it is not possible to estimate the number of journeys for individual tickets. Our methodology disaggregates observations to those with a common flow and common expiry date allowing us to more accurately track small groups of tickets;
- It is acknowledged that not every ticket is read by the gateline. However, we have used a methodology that specifically mitigates this problem. We expect that passengers will always travel on the first day of validity of a weekly Season ticket so we use this with knowledge of the number of sales from LENNON to estimate the gateline capture rate. Capture rates do not necessarily have to be 100% as long as the effectiveness of the gatelines are broadly consistent across the analysis period; and
- Very little data is ready from Smart tickets by a Cubic gateline, therefore we have selected flows with a low proportion of smart card usage to avoid any systematic bias. This might reduce the representativeness of flows with a high proportion of smart data usage.

We selected a total of 15 stations throughout Great Britain (Stage 1, see Chapter 3 for further detail on our station selection). These were selected to:

- provide a geographical spread across the GB network;
- represent both producer and attractor stations (through a mix of Category A and B stations); and
- include a variety of trip distances and fare ratios (i.e. ratio of Weekly Season ticket fare to the walk-up Anytime return fare)

We have also checked, prior to our analysis, whether these stations have active gatelines.

For each station, we selected a total of 24 flows based on the number of LENNON issues (Stage 2, see Chapter 4 for further detail on flow selection), which again covered a comprehensive range of distances, geographies and fare ratios. We have therefore estimated JpT factors representative of the GB network, these have then been averaged with reference to their significance to the whole market to give a GB average JpT value.

The methodology we have developed, as described within this report, has allowed us to estimate JpT estimates for Weekly, Monthly, and Annual Season tickets (Stage 3, see Chapter 5

for further detail on the calculation of the JpT estimates), in what we believe to be more fit for purpose than the existing LENNON assumptions given the available data sources, where the flows with a larger volume of Season ticket travel are weighted to become more significant. We have estimated the JpT figures for each individual flow across the analysis period, inferring the weighted JpT factor for a number of flow types from this.

We have observed a strong correlation between the JpT and the Anytime:Season fare ratio, as well as with distance, which is in line with our intuitive view (Stage 5, see Chapter 7 for further detail on our findings analysis). The JpT figures increase in line with the fare ratio and decrease with distance. This relationship is observed for Weekly, Monthly, and Annual tickets, although it is stronger for the tickets with shorter validity.

We have also analysed the variance of the JpT figures with the flow type; in general, the figures are higher in London than outside London. This result is intuitive and reinforces the conclusions of our analysis.

We have finally estimated a weighted-average JpT figure for each of the three Season ticket validity periods, resulting in the following values[1]:

- A **Weekly** Journeys per Ticket figure of **8.5**; 83% of the current 10.3 LENNON figure;
- A **Monthly** Journeys per Ticket figure of **37**; 82% of the current 45 LENNON figure; and
- An **Annual** Journeys per Ticket figure of **378**, 79% of the current 480 LENNON figure.

The conclusions of our study confirm that the current Season ticket usage indicates a lower JpT figure than currently assumed in LENNON; this conclusion is consistent with trends outlined in other relevant industry studies (see Chapter 1 for reference about other studies). There are, however, some areas of opportunity for further work which might yield improved estimates in some geographies, for instance an assessment based on smart card data where penetration of these products is high.

With the forthcoming changes to LENNON which will potentially allow for journey factor estimates to be specified by their users, our study represents a solid evidence base upon which these new factors could be based. However, the evidence underpinning the use of new factors will need to be reviewed on a case by case basis and may need to be tailored for application to specific purposes.

[1] With the exception of Weekly Season tickets, which have a fixed number of journeys associated with each ticket, the figures shown here for LENNON are an average for a year taken from **http://orr.gov.uk/__data/assets/pdf_file/0014/26600/regional-rail-usage-odm-methodological-report-2017.pdf**

1 Introduction

1.1 Steer and WSP have been jointly commissioned by the Department for Transport (DfT) to deliver this Journeys per Season Ticket Study. The Journeys per Ticket ("JpT") assumptions are key to the calculation (in LENNON) of the number of journeys (and subsequently passenger miles) associated with ticket sales.

1.2 They therefore have an important role in determining estimates of rail passenger demand which are a critical input to much of the modelling and analysis undertaken to produce forecasts of demand and ultimately inform investment decisions. For example, the number of journeys assumed to be undertaken on Season tickets has a bearing on the level of travel anticipated in peak periods which is the time period that tends to drive capacity requirements at stations and on the railway in the form of both network (train paths) and rolling stock capacity.

1.3 The current assumptions are dated and do not align with recent evidence of passenger behaviour e.g. move to part-time and increased home working. The objective of this research is to provide recommendations for updated JpT figures that reflect the current travel behaviours of passengers who use Season tickets.

1.4 In addition, it is proposed to develop these estimates in a way that reflect the different key dimensions that influence the frequency of travel on Season tickets, such as type of flow, distance and ticket price ratio.

Context of the Research

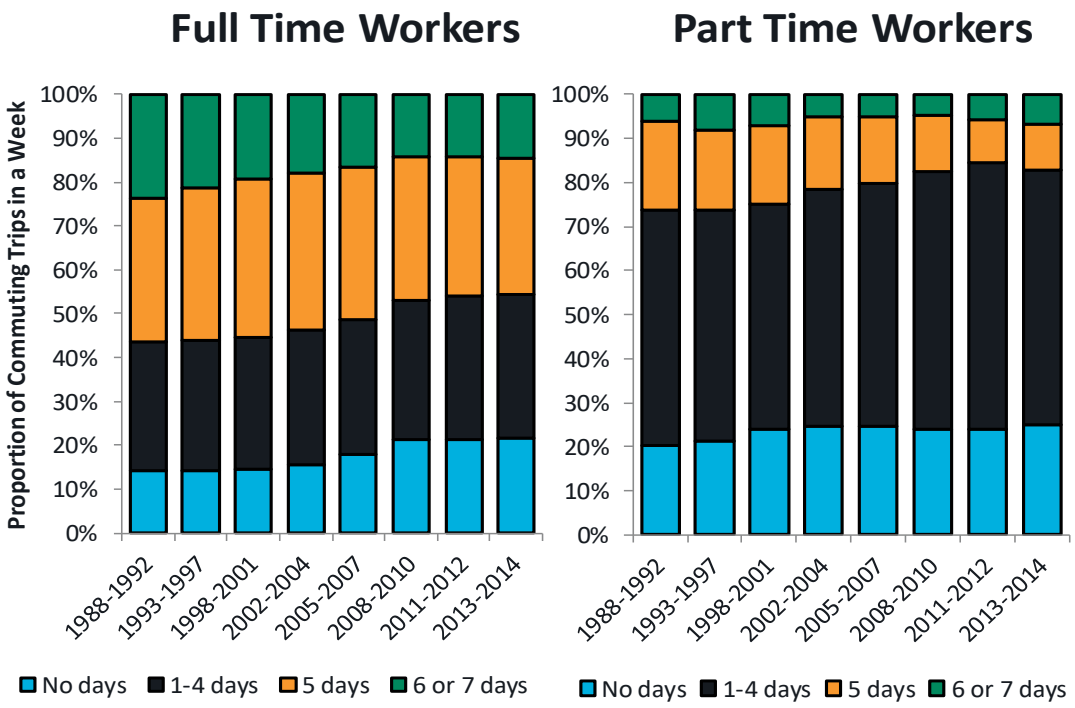
1.5 The travel behaviour of commuters (the primary users of Season tickets) have changed over time and the assumptions that underpinned the number of journeys associated with Season tickets that would have been appropriate in the past may not necessarily be reflective of current and future behaviour and therefore out of date.

1.6 There have been changes to working patterns with increased levels of part-time and home working. As well as the frequency of travel this will also influence the extent to which Season tickets represent value for money compared with other ticket options for that journey.

1.7 Evidence in research and data analysis e.g. analysis of National Travel Survey (NTS) data in the DfT's report on commuting trends,[2] see Figure 1-1 below, indicates that the assumptions currently used are likely to, in general, overstate the frequency of travel for commuting purposes.

[2] Commuting Trends in England 1988-2015, DfT: **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/657839/commuting-in-england-1988-2015.pdf**

Figure 1-1: Proportion of Commuting Journeys in a Week – Full and Part Time Workers, across All Transport Modes



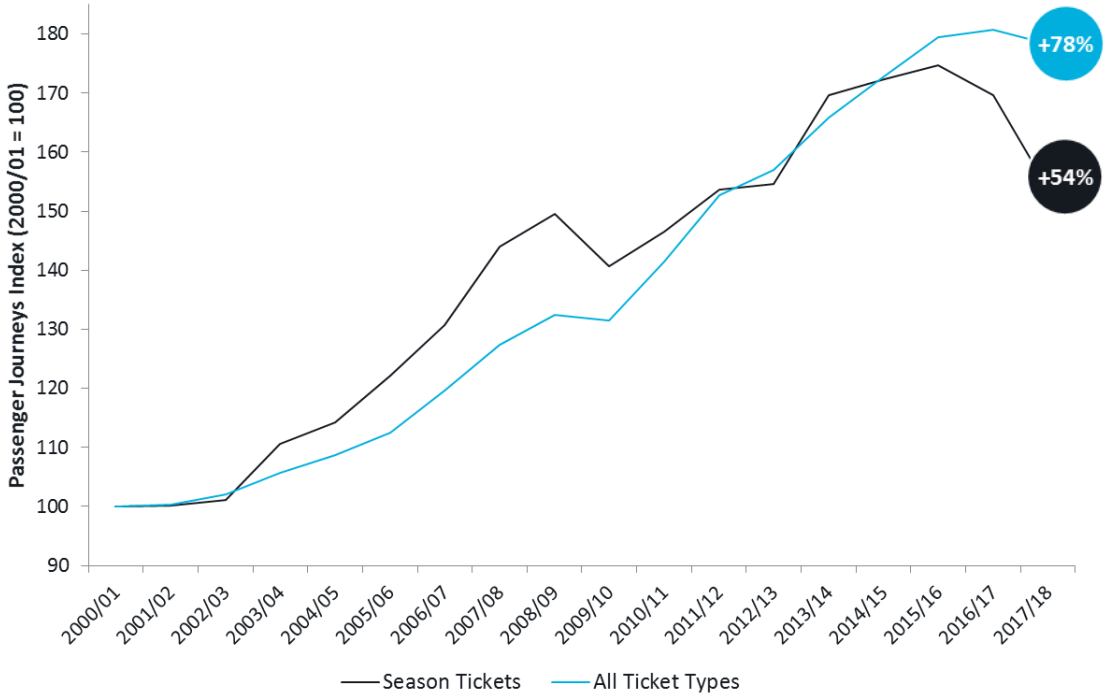
Source: Commuting Trends in England report, informed by the National Travel Survey

1.8 Against this backdrop we see, in recent years, a relative drop in Season ticket journeys whilst total journeys have continued to increase as shown in Figure 1-2 below.

1.9 This suggests that there is some switching of passenger journeys from Seasons tickets to other ticket types which would be reflective of the points made above i.e. as frequency of travel reduces, Full tickets become better value and conversely where 'capping' exists, e.g. with PAYG, alternative (non-Seasons) products become more cost effective for more frequent commuters.

1.10 We would expect this pattern to differ at a more disaggregate level, where frequency of travel differs across types of flow (e.g. by distance of commute).

Figure 1-2: Growth in passenger journeys on Season tickets and All ticket types



Source: ORR Data Portal, Table 12.7 Passenger Journeys by Ticket Type

1.11 Given the changing patterns of commuting this calls into question the appropriateness of the current assumptions that underpin the current Journeys per Ticket estimates used – particularly for Seasons tickets – in rail industry data.

1.12 The Department for Transport have recently undertaken a study on Rail Ticket Use, which shows that the frequency of commuting passengers across the week seems not to be in line with the assumed LENNON factors. For instance, the report found that 44% of commuting passengers travel 5 or more days per week, 30% travel 2 to 4 days per week, and over 25% travel 1 or less times per week.

1.13 Presenting the results by ticket type, this study shows that 85% of Season ticket holders travel 5 or more days per week, 13% travel 2 to 4 times per week and the remaining do so 1 or less times per week[3].

1.14 Presenting the results by ticket type (Figure 26 of the report), this study shows that 85% of Season ticket holders travel 5 or more days per week, 13% 2 to 4 times per week and the remaining 1 or less times.

1.15 However, these results do not show the distance and geography segmentations and might be skewed towards the sampled passengers who are already travelling and are therefore more likely to travel, although overall these seem to confirm the trend outlined by other sources. These assumptions were estimated over 30 years ago and there is no evidence that they

[3] At the time of publication of the Journeys per Season Ticket Study, the report referred to above was unpublished.

necessarily represent the contemporary behaviour of commuters particularly with respect to the two key areas highlighted in the discussion above:

- Commuters still travel on Season tickets but less frequently; or
- Commuters switch from Season to walk-up tickets, where it makes financial sense.

1.16 In the former case, revenue is unaffected (as it is recorded by LENNON) but the actual number of journeys is less than that which is assumed by LENNON, as the factors are not updated to reflect current trends - this leads to an overestimation of the number of journeys associated with these tickets.

1.17 In the latter case, when passengers switch to walk-up tickets, the reported decrease in passenger demand may be higher than in reality, as passengers who travelled on Season tickets were making fewer journeys than assumed and therefore reported and, therefore, the decrease in journeys appears more significant than is likely to be the reality when commuters switch to walk-up fares.

1.18 By undertaking this research, it is hoped to enhance the evidence base to understand contemporary commuter behaviour and provide recommendations for a set of JpT in line with that behaviour.

Structure of Report

1.19 The rest of this report summarises our methodology and key findings and is structured as follows:

- Chapter 2 describes the methodological approach;
- Chapter 3 describes Stage 1 of the methodology – Station Selection;
- Chapter 4 describes Stage 2 of the methodology – Flow Selection;
- Chapter 5 describes Stage 3 of the methodology – Calculation of Journeys per Ticket;
- Chapter 6 describes Stage 4 of the methodology – Validation;
- Chapter 7 describes the Results and Key Findings of the study;
- Chapter 8 describes the quality assurance of the analysis and outputs;
- Chapter 9 describes the key limitations of this study; and
- Chapter 10 presents the conclusions and next steps of the study.

2 Methodological Approach

2.1 This chapter provides an overview of the methodological approach, a description of the data used for this analysis and the key assumptions used in deriving the JpT figures.

Methodology overview

2.2 The underlying basis for the research is to identify a representative sample of flows across Great Britain, categorised by the different factors that are considered to influence frequency of use of Season tickets, and then assess usage on those flows.

2.3 This section presents a high-level description of our methodology. Further detail is provided in the specific chapters covering each of the stages. Our methodology is divided into four stages with a results and finding and final report stages.

Figure 2-1: Stages of our Methodology

Stage 1 – Station Selection

2.4 We have selected fifteen well-managed gatelines that will constitute the basis of our analysis. The gatelines represent a mixture of different geographies, with sufficient volumes so that results of our analysis are representative, and seek that no systematic bias is introduced.

Stage 2 – Flow Selection

2.5 We selected 24 flows per station (noting a small number were eventually discarded), therefore our analysis included a total of 360 flows. We have categorised each flow according to the following features:

- The distance band;
- The PDFH geographical category[4] of the flow; and
- The fare ratio.

2.6 We have sought that there is sufficient variety in flow distance and fare ratio to allow us to analyse and understand the variability of JpT estimates, across these key markets.

[4] This refers to the categories referred to in the Passenger Demand Forecasting Handbook (PDFH) to identify the different geographies and areas of Great Britain.

Stage 3 – Calculation of Journeys per Ticket estimates

2.7 We have identified individual flows and the number of tickets sold on the flow as recorded by LENNON. Weekly Season tickets are identified in gateline data as "MQ" and Monthly and Annual Season tickets as "MT"[5].

2.8 We have first calculated the Weekly, Monthly, and Annual JpT using the methodology described in Chapter 5, which uses the expiry date of gateline and LENNON data to infer our estimates. We have done this for the 360 flows initially selected.

Stage 4 – Validation

2.9 We discovered a small number of flows proved problematic for our process and eventually developed a number of criteria which were used to exclude some of the 360 flows where we viewed the results could not be trusted. These criteria looked at:

- Total number of observations in the gateline data;
- Multiple potential destinations within a ticket (for instance, grouped BR stations); and
- Low capture rate (as the ratio between observations and LENNON data).

Results and Key Findings

2.10 The results of our analysis provide, for each type of Season ticket:

- An average number of journeys for Weekly, Monthly, and Annual Season tickets; and
- An average number of journeys per ticket for the defined distance and PDFH categories.

Reporting

2.11 The final reporting stage, of which this report is one output, summarises all the assumptions, methodology and outputs of our work, alongside the results workbook which has been provided to the Department. This complements the presentation given to the Passenger Demand Forecasting Council on 14 November 2018.

2.12 The ability to implement the methodology is influenced by the availability of data that allows individual Season ticket use to be assessed. In this context, the use of gateline data has been explored as part of the scoping of this research and subsequently incorporated into the methodology implemented.

2.13 The section below reviews the data sources considered and used in the study.

Data Sources

2.14 We have used two primary sources of data to calculate the JpT estimates:

- Gateline data, provided by Cubic; and
- LENNON data.

2.15 Details of these two data sources and an explanation of the methodology used is described in more detail below. We also discuss some other data sources that were considered as part of the methodology development but were not used. Table 2-1 summarises the data used in this study.

[5] A single code 'MT' for Monthly and Annual Season tickets implies that they cannot be distinguished from each other.

Table 2-1: Data sources used in analysis

Data	Source	From	To
Gateline	Cubic	20 March 2017	19 March 2018
LENNON (daily)	RSP (via DfT)	20 March 2017	19 March 2018
LENNON (periodic)	RSP (via DfT)	2015/16, last rail period[6]	2017/18, last rail period

Gateline Data

2.16 Gateline data was requested for a period of one year to allow us to estimate Weekly, Monthly and Annual Journeys per Ticket. LENNON data was requested for the same period, as well as for an additional year prior to the first day for which we have gateline data, as this covers the period when Annual tickets were sold for the passengers travelling within the core year of our analysis.

2.17 The use of the gateline dataset has its strengths and weaknesses, outlined in Table 2-2. The main strength is the sheer wealth of the dataset and the capturing of ticket type, although it is to note that this is limited to only the two-letter product code being captured by the gateline rather than the full four-letter CTOT (class and type of ticket).

2.18 Note that the class of ticket, denoted by a 1 (First Class) or 2 (Standard Class) at the start of the CTOT, is provided as a separate field in the gateline data but was not required for the analysis.

Table 2-2: Strengths and Weaknesses of Gateline Data

Strengths	Weaknesses
• Wealth of data • Captures ticket type[7] • Visibility of day used • Visibility of expiry date • Continuous timeline • Large number of gated stations	• No unique identifier • Gateline leakage • Smart tickets

2.19 Our methodology has been developed to mitigate these weaknesses of the gateline data, specifically:

- **No unique identifier:** We have disaggregated to tickets with a common expiry date to estimate their journeys per ticket. The data does not allow us to trace through the journey of a particular ticket, and allows us to instead investigate factors for small subsets of the data;
- **Gateline leakage:** As described on the Section describing the calculation of the Journeys per Ticket, some leakage is not an issue for our methodology, as along as the gateline behaviour is consistent over time; and
- **Smart tickets:** Cubic gateline data does not include sufficient data for analysis of smart tickets, therefore we have selected stations and flows as part of our methodology without a significant proportion of travel on smart tickets, in order to avoid any systematic bias. This

[6] Rail ticket sales data is sorted into periods and there are thirteen periods of 27 to 28 days per financial year.

[7] As described in more detail in the section relating to the calculation of the Journeys per Ticket figures, the ticket type is only presented as a two-digit code, therefore it does not always allow to distinguish between ticket types, such as between Monthly and Annual tickets.

might, however, reduce the representativeness of flows with a high proportion of smart data usage.

LENNON data

2.20 We requested four LENNON queries for our analysis, which respond to the following specifications. Where any of the gatelines form part of a station group (for example London Marylebone in the London group, or Manchester Piccadilly in the Manchester group), the corresponding "BR" origin/destination code was also included in the filter.

Table 2-3: LENNON Query 1

Filter Description	Filter Operation	Filter Input
Date of Settlement	between	20/03/2017,19/03/2018
Origin Code	=	Selected NLCs
Product Level 1 Code	=	PG04 and PG08 (Seasons)

Table 2-4: LENNON Query 2

Filter Description	Filter Operation	Filter Input
Date of Settlement	between	20/03/2017,19/03/2018
Destination Code	=	Selected NLCs
Product Level 1 Code	=	PG04 and PG08 (Seasons)

Table 2-5: Reported Fields in LENNON Queries 1 and 2

Reported Fields
Date of Travel
Season Ticket End Date
Origin Code
Destination Code
Product Level 2 Code
Product Level 2 Description
Issues

Table 2-6: LENNON Query 3

Filter Description	Filter Operation	Filter Input
Period of Settlement	between	2016/P13, 2018/P13
Origin Code	=	Selected NLCs
Product Level 1 Code	=	PG04 and PG08

Table 2-7: LENNON Query 4

Filter Description	Filter Operation	Filter Input
Period of Settlement	between	2016/P13, 2018/P13
Destination Code	=	Selected NLCs
Product Code	=	PG04 and PG08

Table 2-8: Reported Fields in LENNON Queries 3 and 4

Reported Fields
Period of Settlement
Origin Code
Destination Code
Product Level 1 Code
Product Level 1 Description
Issues

Other Data Sources

2.21 LENNON data was initially considered for the derivation of fares for the selected flows, to produce the analysis relating journeys per ticket to fare ratios (Anytime Return:Weekly Season).

2.22 Concerns over the specification of this data, specifically being able to obtain the appropriate adult, undiscounted fare for each flow and relevant ticket type, meant this option was discarded in favour of using publicly available fares data from brfares.com.

2.23 For each flow in the analysis, the weekly Season and Anytime Return fares were obtained. In each case, where more than one route was available for the flow, the Any Permitted fare was taken, and where both an Anytime Return and Anytime Day Return were available, the Day

Return was prioritised. The project team viewed that this better represents the trade-off between Return and Season tickets, which will both typically involve daily return journyes.

2.24 In addition, other data sources were considered but discarded, as they were not suitable to be used as part of our approach. The following paragraphs justify our decisions.

Passenger Counts

2.25 The option of using passenger counts to validate the capture rate was considered by the project team. In conjunction with gateline data, this could provide a view of the proportion of passengers not validating their tickets at the gateline, which would give assurance on the robustness of the JpT estimates calculated using the gateline data.

2.26 However, we note the following potential issues with the passenger counts data:

- **Non-overlapping data**: The most recently-available counts data available at the time of the analysis was for Autumn 2016, which did not align with the date range for the gateline data;
- **Mismatch of Information**: For many stations, the count data is calculated using Automatic Passenger Counts, and is therefore an average across a number of days, and not necessarily the same days for each service. Individual daily comparison between gateline data and passenger counts for validation purposes is therefore difficult and almost impossible; and
- **Incompatible Level of Disaggregation**: Passenger counts are not disaggregated by flow or ticket type, and therefore can only validate total gateline observations. Our analysis data is limited to gateline observations on very specific flows, and also includes Season tickets only.

2.27 Due to the degree of uncertainty introduced by applying an uplift factor to Autumn 2016 counts (if we were to estimate Autumn 2017 counts) and the lack of suitable disaggregation, we believe LENNON data is our best option for the purpose of understanding the gatelines' capture rates.

2.28 The focus of this study is to calculate new JpT estimates for Season tickets. To do this, we have compared LENNON Season ticket transactions with gateline observations of Season tickets on their first day of validity. This has allowed us to specifically validate the captured proportion by flow and ticket type.

Smart Ticket Data

2.29 Smart ticket data was not available for use on this project. Our understanding is that the data can provide unique identifiers for Smart Card holders, allowing for the calculation of frequency of travel for individual Season ticket holders. In return, using such disaggregated data might cause an issue with regards to the new data protection regulation (GDPR), so this would need to be taken into account.

2.30 This would provide greater disaggregation than the methodology currently in place, which aggregates some passengers together where their Season tickets have the same expiry date. However, such rich data does not yet exist for the entire network due to the low uptake of Smart Tickets with limited geographical areas covered.

2.31 The objective of this study is to provide JpT estimates with a greater level of aggregation. We therefore opted to concentrate on conventional paper tickets that better represent the demand in scope.

Network Rail Station Data

2.32 We have considered the use of Network Rail automated CCTV count data. Manual counts using CCTV data were used in other studies to supplement the gateline data (i.e. in a recent study

undertaken by the project team for Chiltern); however, whether done manually or using automatic counts, this data would not capture flow information or ticket type, and therefore would not provide any additional confidence in the calculated JpT estimates.

Summary of Assumptions

2.33 Table 2-9 summarises the key assumptions of our methodology:

Table 2-9: Summary of our assumptions in the analysis of Journeys per Ticket estimates

No.	Category	Description
1	Capture Rate	We initially assumed that 100% of Weekly Season ticket holders make a journey on the first day of ticket validity as a way to understand the maximum capture rate of a gateline. Our results, however, indicate that this is not always the case, particularly when tickets are purchased over weekends. We have revised our assumption and now assume the number of tickets observed by the gateline is the greater of either the first day of validity or 90th percentile of throughout the period of validity. This, divided by the total number of tickets in circulation from LENNON, determines the capture rate. Further clarification about this is provided in Section 6.
2	Capture Rate	We have assumed that the observed capture rates for Weekly tickets will also apply to Monthly and Annual tickets.
3	Return Travel	We have analysed gateline data for one direction only, and therefore assumed that all passengers on Season tickets will also return by rail.
4	Weekends	We have assumed that low capture rates at weekends are due to low use of tickets, as opposed to gates being left open. This has been validated through analysis of the Marylebone gateline.

3 Stage 1 – Station Selection

3.1 The first stage of our analysis was to select well-managed gatelines, representing a mixture of different geographies and flows across the rail network. We selected gatelines with sufficient volumes resulting in results of our analysis being robust, while seeking to prevent the introduction of systematic bias due to our selection/sampling.

3.2 The process for selecting the stations and the final list used to calculate the JpT estimates is described in this chapter.

Selection Criteria

3.3 To analyse the different drivers of the number of journeys per ticket, we have considered several selection characteristics for the selection of stations ensuring that all variants are represented in our analysis. The characteristics we considered were:

- **Geography**: observations from across the rail network including consideration of whether the station includes London and non-London flows;
- **Distance**: classified qualitatively as a first step as to whether the primary markets to/from that station are likely to be short, mid-range, or long-distance flows. It is noted that most stations serve a combination of short, mid-range and long-distance flows and this categorisation was only used to initially obtain a fair spread of different stations by typical usage; and
- **Station category**: we have selected stations in categories A and B (the top two categories that have a higher footfall), so that the analysis is sufficiently representative, but not necessarily biased towards category A stations (noting too that many smaller stations will not have gatelines).

3.4 We have considered that the range of shortlisted stations would contain flows with sufficient variability in ticket fare ratios that allow us to also analyse the JpT estimates against this variable.

3.5 In addition to these selection criteria, the validation process (Stage 5) has allowed us to determine whether the capture rate for the gatelines is sufficiently high to produce robust Journeys per Ticket estimates, with those with a high leakage rate being potentially removed from our list.

3.6 The railway geography can be divided in ten Primary National Location Code (NLC) zones, to which every railway station and ticket issuing point can be uniquely allocated. Therefore, we have aimed, where possible, to spread our selected stations across these zones to obtain as even a representation as possible. Table 3-1 below lists the different NLC zones.

Table 3-1: NLC Zones

NLC Zone	Geographical Area
1000-1999	West Coast Main Line (southern section); West Midlands (north of Birmingham); Midland Main Line (southern section); London Marylebone-Aylesbury section of Chiltern line; Cumbrian Coast Line
2000-2999	West Coast Main Line (northern section); Merseyside; Greater Manchester; North Wales Coast Line
3000-3999	Great Western Main Line and associated branches; Chiltern main line; South Wales and Welsh/English border area
4000-4999	Lines in West of England and Wales north of those in "3000" series; ex-Great Western Railway lines in West Midlands
5000-5999	All lines south, south-east and south-west of London
6000-6999	East Coast Main Line (southern section); Great Eastern Main Line; lines in eastern England; suburban lines north and east of London; Midland Main Line (northern section)
7000-7999	East Anglia; East Coast Main Line (northern section); associated branch lines in North-east England
8000-8999	North and North-east England east of the Pennines, including most of Yorkshire; Highlands and north of Scotland
9000-9999	Central and Southern Scotland, including suburban Glasgow and Edinburgh

Station Selection

3.7 Based on the selection criteria set out above, we selected an initial fifteen stations which were then shared with the Passenger Demand Forecasting Council (PDFC) for comment and understand their views on both our approach to this study and our initial shortlist of stations. These stations were:

- London Marylebone;
- Cambridge;
- Reading;
- Newcastle;
- Brighton;
- Coventry;
- Nottingham;
- Glasgow Central;
- Cardiff Central;
- London Charing Cross;
- Basingstoke;
- Swansea;
- Huddersfield;
- Peterborough; and
- Chester.

3.8 We received feedback from a number of stakeholders, which can be found in Appendix A. Based on this feedback we replaced some of the stations initially included in our selection as per the reasons set out:

- **Nottingham**: the feedback received suggested that only one of the two entrances/exits is gated, so Nottingham was excluded and an alternative station, Derby, in the same geographical area included; and
- **Charing Cross**: the feedback received indicated that for tickets with multiple London termini choices ('London Terminals'), it's generally not possible to infer the volume of passengers using any particular termini station from the LENNON data, so the validation process will not be possible. We replaced Charing Cross with Liverpool Street which, while also having London Terminals tickets, would only have a few interchange points, e.g. Stratford or Tottenham Hale, and we consider that it would not be financially sensible for passengers to buy a London Terminals ticket and then interchange in Stratford as opposed to buying a Stratford ticket and then using the tube; therefore, we consider that Liverpool Street would capture most of the journeys on London Terminals tickets[8].

3.9 In addition, some of the stations in our original station selection were not Cubic-owned gatelines or Cubic were not able to provide data – this required us to revise the choice of some stations and include alternatives where Cubic held the gateline data. Here we selected stations with similar features to those originally selected, both in terms of geography and types of flows likely to be using the station.

3.10 The following stations, originally selected, were replaced by other stations with Cubic gatelines:

- Newcastle, replaced with Crewe;
- Nottingham, replaced with Derby; and
- Peterborough, replaced with Milton Keynes Central.

3.11 The final list of 15 stations selected are shown in Table 3-2 below.

[8] Note we do not need 100% of demand to be via Liverpool Street, just a large enough proportion – if a proportion of demand uses an alternative route and evades our study gateline this will be accounted for in the same way as other gateline leakage.

Table 3-2: Final Station Selected

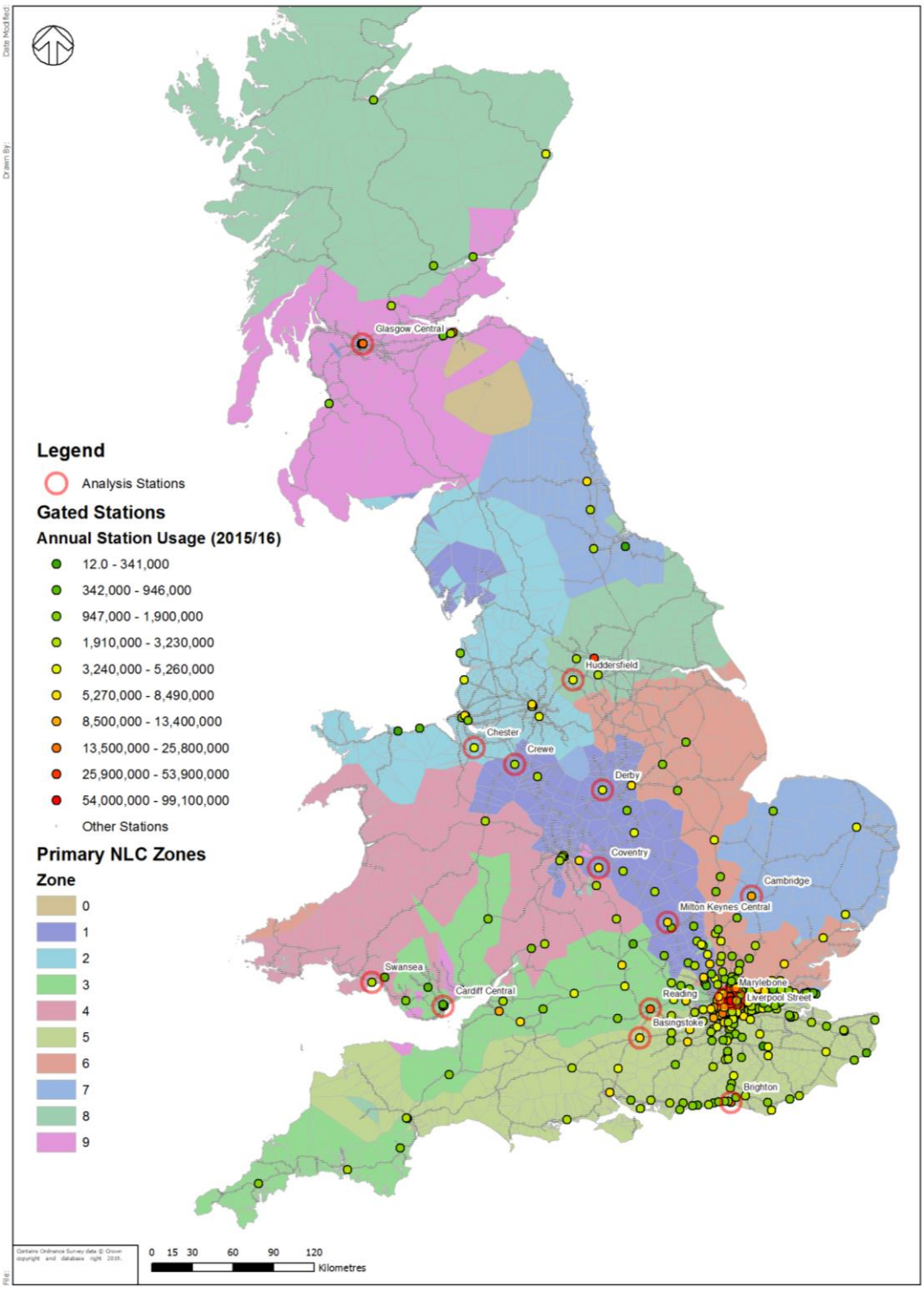
Station	NLC	No. Gates	TOCs[9] (SFO in bold)	Station Category	Geography		Distance		
					London	Non-London	Short	Mid-Range	Long
Marylebone	**1475**	19	**CH**	A	✓	✗	✓	✓	✓
Cambridge	**7022**	15	**GA**, GN, and XC	B	✓	✓	✓	✓	✗
Reading	**3149**	39	GW, SW, and XC (**NR**)	B	✓	✓	✓	✓	✗
Crewe	**1243**	9	AWT, WMT, EMT, NT, XC, **ICWC**	B	✗	✓	✓	✓	✗
Brighton	**5268**	26	**SN**, TL, and GX	B	✓	✓	✓	✓	✗
Coventry	**1030**	11	**ICWC**, WMT, and XC	B	✗	✓	✓	✓	✗
Derby	**1823**	10	XC, **EMT**, NT	B	✓	✓	✓	✓	✗
Glasgow Central	**9813**	63	ICWC, ICEC, XC, SR, and TPE (**NR**)	A	✗	✓	✓	✓	✓
Cardiff Central	**3899**	13	**ATW**, GW, and XC	A	✗	✓	✓	✓	✓
Liverpool Street	**6965**	66	GA (**NR**)	A	✓	✗	✓	✓	✓
Basingstoke	**5520**	11	**SW**, XC, and GW	B	✗	✓	✓	✓	✗
Swansea	**4222**	6	**ATW**, and GW	C1	✗	✓	✓	✓	✗
Huddersfield	**8437**	9	**TPE**, and NT	B	✗	✓	✓	✓	✗
Milton Keynes Central	**1378**	7	**WMT**, ICWC, SN	B	✓	✓	✓	✓	✗
Chester	**2412**	5	**ATW**, NT, ME, and ICWC	B	✗	✓	✓	✓	✗

3.12 This set of stations covers a wide range of geographies, flow distances, and station categories, therefore we consider it is sufficiently representative for the JpT analysis.

[9] Glossary of TOCs:
Chiltern (CH), Greater Anglia (GA), Great Northern (GN), CrossCountry (XC), Great Western (GW), South Western (SW), InterCity East Coast (ICEC), InterCity West Coast (ICWC), Northern (NT), TransPennine Express (TPE), South Eastern (SE), Southern (SN), Thameslink (TL), Gatwick Express (GX), West Midlands Trains (WMT), East Midlands Trains (EMT), ScotRail (SR), Merseyrail (ME), and Arriva Trains Wales (ATW); Network Rail (NR).

3.13 An overlay of the selected stations over the different NLC zones is shown in Table 3-1. The colour of the station represents the station total entries and exits (based on ORR Estimates of Station Usage), with a darker colour representing a higher number of total entries and exits, and a red circle representing our selected stations.

Figure 3-1: Gated stations in Great Britain and NLC zones including selected stations for analysis



3.14 In addition, we have analysed the percentage of Season ticket users in accordance with the statistics published by the ORR (2016/17 Estimates of Station Usage), as shown in Table 3-3.

Table 3-3: Season ticket entries and exits by station (2016/17 ORR Estimates of Station Usage)

	16/17 Entries & Exits		
Station	All tickets	Season tickets	% Season
Marylebone	16,666,936	5,750,164	34.5%
Cambridge	11,424,902	3,441,060	30.1%
Reading	17,122,000	5,949,720	34.7%
Crewe	3,085,604	447,978	14.5%
Brighton	15,993,072	4,244,710	26.5%
Coventry	7,377,584	1,984,164	26.9%
Derby	3,962,860	585,270	14.8%
Glasgow Central	32,060,134	9,991,858	31.2%
Cardiff Central	12,534,884	3,454,506	27.6%
Liverpool Street	67,339,218	33,420,594	49.6%
Basingstoke	4,774,744	1,122,564	23.5%
Swansea	2,130,154	386,092	18.1%
Huddersfield	5,092,542	1,540,042	30.2%
Milton Keynes Central	6,851,324	2,231,786	32.6%
Chester	4,649,800	507,602	10.9%

3.15 All the selected stations have a significant number of Season ticket users therefore, provided the gateline data has a sufficiently high capture rate at these locations, we consider they should be appropriate for this analysis.

4 Stage 2 – Flow Selection

4.1 The second stage of our analysis was to identify a comprehensive sample of flows involving the selected stations on which to undertake the analysis. The process for selecting the flows is described in this chapter.

Selection Process and Criteria

4.2 . A LENNON dataset was obtained including all Season ticket flows to/from the stations which were selected at Stage 1. From these stations the biggest flows in terms of Season ticket issues have been identified.

4.3 Table 4-1 and Table 4-2 present the LENNON filters and fields that were reported within the dataset.

Table 4-1: LENNON filters and operation, as specified on flow selection

Filter Description	Filter Operation	Filter Input
Date of Settlement	between	20/03/2017,19/03/2018
Origin Code (Query 1) /Destination Code (Query 2)	=	0403,0433,1030,1072,1243,1378,1475,1823,1947,2412,2771,3149,3899,4222,5268,5520,6133,6965,7022,7728,8437,8487,9922
Product Level 1 Code	=	PG08

Table 4-2: LENNON fields specified on flow selection

Reported Fields
Date of Travel
Season Ticket End Date
Origin Code
Destination Code
Product Level 2 Code
Product Level 2 Description
Issues

4.4 This process for identification of key flows was successful for all stations except Marylebone and Liverpool Street. The majority of Season tickets to these stations are issued with a destination of London Terminals, with the specific London termini only appearing when the ticket is bought at the London station.

4.5 Selection based on the latter was viewed as introducing bias, given it would be reflective of the top attractor locations (other than London) on the route, whereas the most relevant flows are those where these stations are the producer.

4.6 These tickets represent only a small proportion of the total Season ticket demand for these stations, and so would be liable to miss some of the larger producer stations (for example Warwick Parkway).

4.7 For these stations (Liverpool Street and Marylebone), we used one weeks' worth of gateline data to obtain the 30 flows with the highest number of gateline observations, as opposed to obtaining the flows using LENNON data directly. The top 24 of these flows selected.

Flow Categorisation

4.8 We have categorised flows by:

- Distance;
- Fare Ratio; and
- PDFH Flow Type.

4.9 Distances have been obtained from the DfT's RUDD[10] dataset. For any flows not contained in RUDD, we have sourced the distance from *railmiles.me*. These distances have been banded, for the purposes of cross-tabulation of the results, into flows:

- Less than 20 miles;
- 20-100 miles; and
- Over 100 miles.

4.10 The above distance categories have been selected to be in line with the typical distance bands used in PDFH, although we are aware that there could be a certain degree of variability within each of these categories (particularly those between 20-100 miles).

4.11 Fare ratios have been calculated using data from *brfares.com*. The ratio has been calculated as the relativity between the 7-day Season and the Anytime (Day) Return fare for each flow, as available from the May 2018 fares round. Where both an Anytime Day Return and an Anytime Open Return are available on a flow, the Anytime Day Return fare has been used for the ratio.

4.12 Although these fares do not directly correspond to the period used for the analysis of gateline data, it is unlikely that the ratio of fares has changed significantly between 2017 and 2018, particularly given both are regulated on most commuter flows.

4.13 The approach to allocate flows to PDFH flow types is displayed in Table 4-3.

[10] Rail Usage Demand Drivers (RUDD). RUDD is a comprehensive dataset of rail stations and flows, which includes a range of socioeconomic and rail specific variables measured over time. In particular, it includes flow distances, which we have used for this analysis.

Table 4-3: Mapping of Flows to Flow Types

	Destination is London	Destination is a city in the South East	Destination is a City	Destination is in the South East	Destination is Rural
Origin is London	London South East	London South East	London Rest of Country	London South East	London Rest of Country
Origin is a city in the South East	London South East	South East Inter-Urban	Non-South East Inter-Urban	South East Non-London	South East Non-London
Origin is a City	London Rest of Country	Non-South East Inter-Urban	Non-South East Inter-Urban	Non-London	Non-London
Origin is in the South East	London South East	South East Non-London	Non-London	South East Non-London	Non-London
Origin is Rural	London Rest of Country	South East Non-London	Non-London	Non-London	Non-London

4.14 This gives 6 different flow types to which the JpT estimates have been aggregated:

- London South East;
- London Rest of the Country;
- South East Inter-Urban;
- Non-South East Inter-Urban;
- South East Non-London; and
- Non-London: Remaining flows that do not fit within the other above categories.

Summary of Flows

4.15 In total we have included 360 flows across 15 stations in GB, and the geographical spread of these flows can be seen in Figure 4-1. There is one equally-sized shaded dot per flow used providing a shading or 'coverage' to spatially visualise the commuter regions and lines-of-route we have incorporated. The map indicates that a broad range of commuter distances and different regions in GB have fed into the analysis.

4.16 The original scope proposition for this research indicated analyse a minimum of 200 flows would be required for an adequate sample size, therefore even after our validation and exclusion process described in Stage 4, these 360 flows are reduced to a number well over 200 in total.

4.17 We have also generated maps for each of the stations individually as part of our analysis and investigation of the flows we have incorporated, indicating the journeys per ticket for each flow in a colour-bar style.

4.18 Appendix B contains maps of all 360 flows across all the 15 stations used in the analysis.

Figure 4-1: Map of all stations and flows used in the analysis

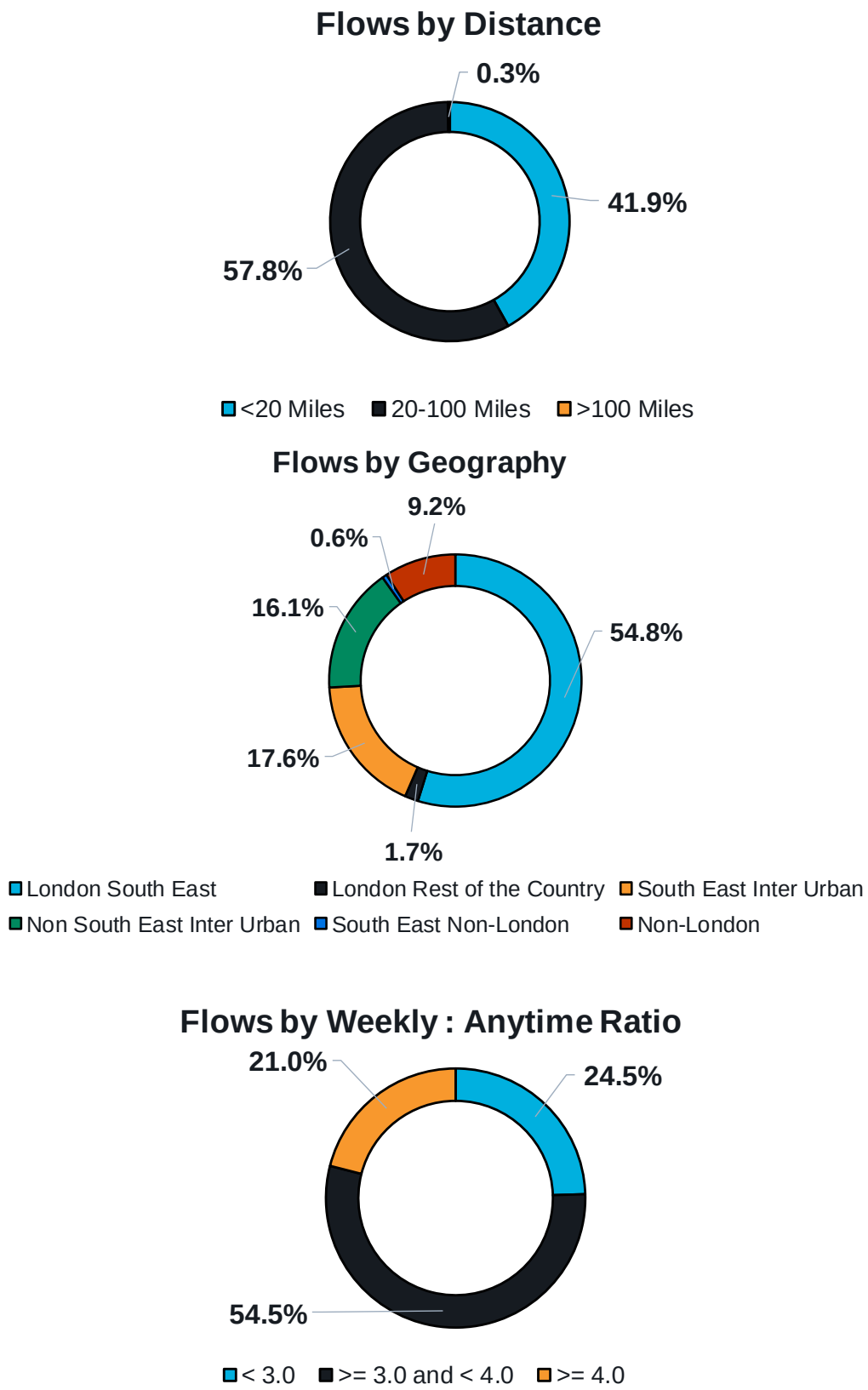
4.19 Table 4.4 presents the total number of flows selected for each of the distance and flow type categories, prior to the validation exclusions undertaken, which are presented in Chapter 6.

Table 4.4: Flows selected by distance and flow type category, prior to validation

Flow Type Category	Under 20 miles	Between 20-100 miles	Over 100 miles
London South East	12	51	0
London Rest of the Country	0	5	4
South East Inter Urban	62	19	0
Non South East Inter Urban	52	50	5
South East Non-London	2	8	0
Non-London	61	29	0
All Flow Types	189	162	9

4.20 The graphs below in Figure 4-2 show the proportion of flows in each of the distance, geography and fare ratio categories.

Figure 4-2: Proportion of flows by category



4.21 The proportion of flows in each of the categories provides a broad representation of Season ticket flows across Great Britain. We have weighted our results, calculated in Section 7, by the proportion of Season ticket demand in each category (sourced from RUDD), following the breakdown in Table 4-5 below.

Table 4-5: RUDD Proportions by Category

PDFH Category	<20 Miles	20-100 Miles	>100 Miles
London South East	41.3%	31.2%	0.1%
London Rest of the Country	< 0.1%	1.6%	0.5%
South East Inter Urban	7.7%	1.2%	< 0.1%
Non South East Inter Urban	5.2%	2.3%	< 0.1%
South East Non-London	0.1%	0.3%	< 0.1%
Non-London	5.5%	2.9%	< 0.1%

4.22 This has allowed us to estimate an overall average Journeys per Ticket value representative of all Season ticket flows in Great Britain.

5 Stage 3 – Calculation of Journeys per Ticket estimates

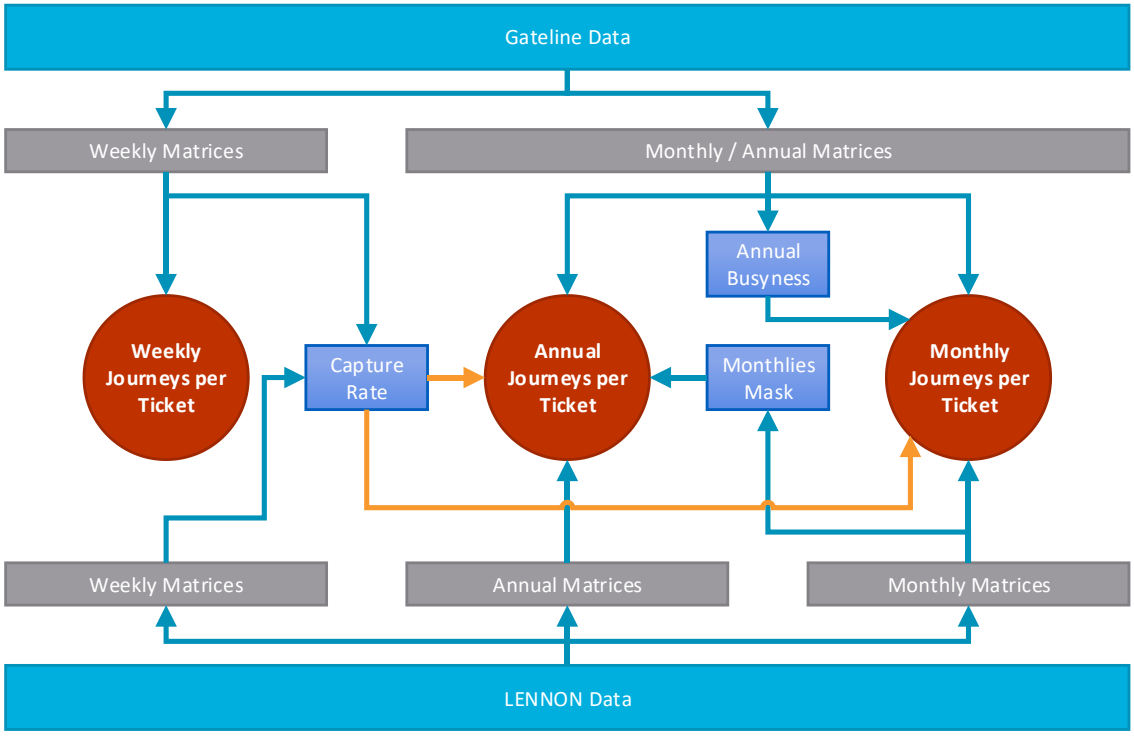
Approach

5.1 Our approach has been designed around harnessing the strengths of gateline data while addressing and mitigating its limitations. This includes developing a robust process for determining a capture rate due to gateline leakage.

5.2 As described in Section 2, this data has no unique ticket identifier and we have therefore disaggregated the data by expiry date for the range of dates where the tickets were used. Expiry dates from the first day of our analysis year, to one year after the final day were included in our analysis.

5.3 The flowchart in Figure 5-1 outlines the process and the following subsections will explain some of the stages (in the blue boxes) to get to the JpT values for each Season ticket type (red circles).

Figure 5-1: Flowchart of overall approach for calculating Journeys per Ticket from gateline and LENNON data



5.4 In addition to gateline leakage, a key characteristic of the data which drives aspects of the analysis methodology is that the gateline cannot distinguish between Monthly and Annual tickets as they are coded identically. An assessment of 'Annual busyness' and a 'Monthlies mask' is therefore used, both of these are explained in more detail in the subsections below.

5.5 The analysis has been undertaken in a series of Excel workbooks which are being provided to the DfT alongside this report.

Methodology

5.6 We have analysed Season ticket gateline observations using the following information:

- Flow;
- Direction of Travel (Gate in Entry or Exit mode);
- Ticket Validity;
- Ticket Type;
- Date of Use (when the ticket passed through the gateline); and
- Ticket Expiry Date.

5.7 We have produced one workbook for each of the selected 15 stations configured to capture the 24 top flows by volume. This total (360 flows) provides a much larger sample than originally proposed (200) so that there is considerable contingency available to eliminate flows should we discover anomalies or discrepancies on a flow which we view might make the results unreliable.

5.8 Tickets have been considered as part of this flow if the Origin NLC and Destination NLC match. The directionality has not deemed to be relevant since we considered that there may be some passengers who purchase Season tickets in the 'homebound' direction and doing so would not change or limit their ability to travel.

5.9 For this study, only journeys arriving at the respective station were included. This is considered slightly cleaner since we will only capture each individual once per round trip, allowing us to understand the travel behaviour but cut the number of searches required in half. Therefore, the gateline records used were those with the gate 'in exit mode'. We only include observations where the ticket was accepted as 'valid' by the gateline.

5.10 The ticket type was limited to either Weekly Season tickets or to Monthly and Annual Season tickets. Tickets fed through the gateline are assigned a different ticket type code depending on ticket type: Weekly Season tickets are identified as "MQ", Monthly+[11] and Annual Season tickets as "MT" (without the possibility to distinguish between both) and "ZT" as replacement tickets. For this reason, Monthly and Annual tickets are shown together in the results below.

Data Miner Tool

5.11 Our Data Miner tool processes the gateline data line by line and extracts the relevant ticket information. Below is a simplified summary of how the relevant information is extracted:

1. The user sets up the initial conditions that the program will use to search through the data, i.e. the time period for the search, the ticket type being searched for, the direction of travel through the gateline that will be counted and a list of the flows which are to be included
2. The program opens the first file in the specified list of gateline csv data files and begins reading the gateline data;
3. The expiry date, the date of use, and the flow direction information on the ticket is saved if the record matches the relevant parameters;
4. The program increments the number matching the same expiry date, the date of use, and the flow. This information is saved as a three-dimensional array;

[11] Strictly monthly and custom tickets with a duration over a month

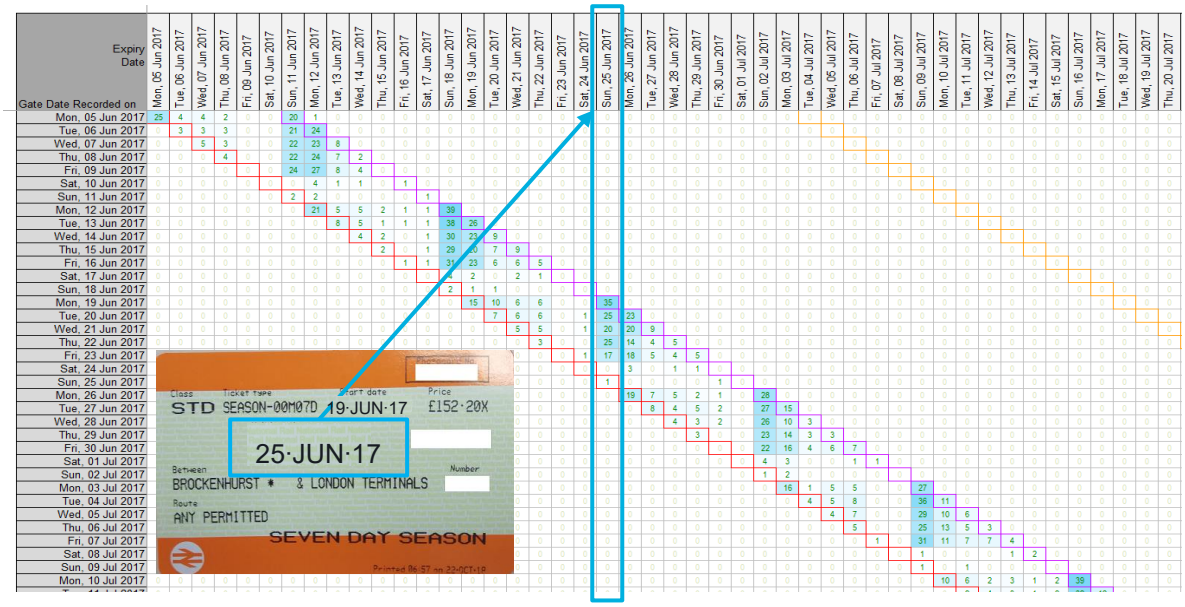
5. The process in 3. and 4. is looped until the program has read the entire file. If there is another gateline data file in the list to read, the program opens this file and repeats steps 3. and 4. If there is no other file to read, the program moves onto writing the data; and
6. The output data is produced, writing a two-dimensional matrix representing each flow on a separate tab in an Excel workbook: the two dimensions are date of the observation (down the worksheet) and the expiry date (across the columns). The data within these arrays show the number of tickets, for each flow, with the same expiry date that were used on each day of the study period. An example of the output layout can be observed in Figure 5-2.

5.12 Depending on station size, the tool can take anything from half an hour (London Marylebone) to four hours (London Liverpool Street) to run.

Weekly Journeys per Ticket

5.13 From the processed gateline data we have obtained a matrix like the one shown in Figure 5-2 below. If we isolate a single column, we can view the journeys observed on Season tickets which expire on the date indicated by the column header; if we consider one of these columns (e.g. Sunday 25th June highlighted below), we can see how many people with these tickets were observed at the gateline each day for that week.

Figure 5-2: Weekly Season Tickets Observations



5.14 As we might expect, the busiest day in the example above is the first day of validity, Monday 19th June, with 35 passengers observed. For subsequent weekdays we observe 25 passengers on Tuesday, 20 on Wednesday and 25 on Thursday. On Friday there are 17 observations, which is a noticeable decrease, and then the number of observations reduces to no observations on Saturday and one on Sunday.

Calculation Steps

1. To calculate the journeys per Weekly Season ticket we have assumed that all passengers who hold a Weekly ticket travel on its first day of validity. For example, for each column of the observations vs expiry date matrix, shown in Figure 5-2, we can find the number of observed journeys on the first day of validity, and then the sum of all journeys in this column

of the matrix. This sum represents all observations of journeys exiting the gateline at a given station.

As described in Section 1, to circumvent occasions where the first day of validity appears to underrepresent the number of tickets in circulation, we have also calculated the 90th percentile of observations across the week for the given expiry date, and have taken the larger of the two values to be representative of the tickets in circulation (see a slight refinement of this calculation in paragraphs 5.17 to 5.20)

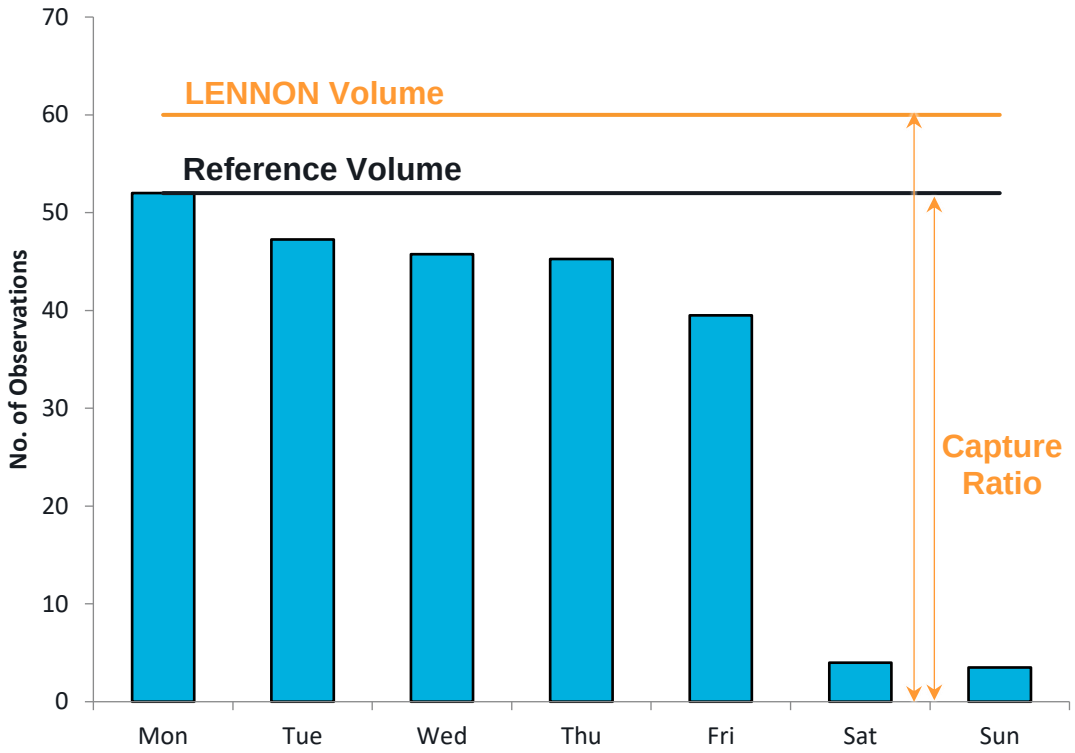
2. As we are assuming that all journeys on Season tickets will be return journeys, we multiply the sum of observations by two to get the respective number of journeys. Finally, dividing the sum by the number of first day observations gives us the number of journeys per ticket. For example (displayed in Figure 5-3), on a Monday we find that there are 52 observations on the first day of travel, and we therefore assume, for the purpose of the calculation, that the capturable reference volume is 52 tickets.

   We know that the corresponding volume of tickets in circulation, as recorded in LENNON, will be higher than the figure calculated above. Assuming that variations in gateline capture rates across the study period are evened out, we use the ratio between the inferred reference volume and the LENNON volume to calculate the gateline capture rate. In the below example, we assume 60 tickets are in circulation and therefore the implied capture rate is 87%.

3. In the remainder of the column we find that there are 237 journeys in total from Monday to Sunday. We multiply 237 by 2 to get the observed number of journeys, and divide this by the 52 observations reference volume find the journeys per Weekly ticket that have the expiry date of the Sunday. In this example the JpT value is 9.1.

4. This calculation is performed for every column of the matrix to find a journey per ticket for each ticket expiry day.

Figure 5-3: Weekly Season Ticket Journey per Ticket Calculation



Source: Illustrative example to present approach graphically

5.15 As it is acknowledged that the gateline does not capture every ticket that travels, data regarding the real number of valid tickets in existence, provided from LENNON, is used to calculate the ratio of tickets observed to the total number of valid tickets (the "capture ratio").

5.16 Again, we use the assumption that all tickets on their first day of validity are used, and from this we compare the number of tickets observed to the total number of valid tickets. In Figure 6-3 we show the LENNON volume compared to the Reference, or observed volume, on the first day of validity which in this example is Monday.

5.17 In a small number of cases, particularly where the first day of validity was a weekend, we observed that travel on subsequent days was a higher than the first day. To accommodate these days, we have included another method for calculating the reference volume called the 90th percentile method.

5.18 The 90th percentile method (as the name suggests) calculates the 90th percentile value of all observations of tickets with the given expiry date, and compares this to the number observed on the first day of validity. If the 90th percentile value is larger than the number of tickets observed on the first day of validity, it is assumed that the 90th percentile value is a more plausible estimate of the reference volume for a given expiry date.

5.19 In this case, we calculate the capture ratio by comparing the 90th percentile value to the number of valid tickets, rather than using the observations from the first day of validity.

5.20 We did investigate working without this catch, however the results found were much noisier and after reviewing both sets of results we viewed the set including the 90th percentile catch to be the more reliable answer.

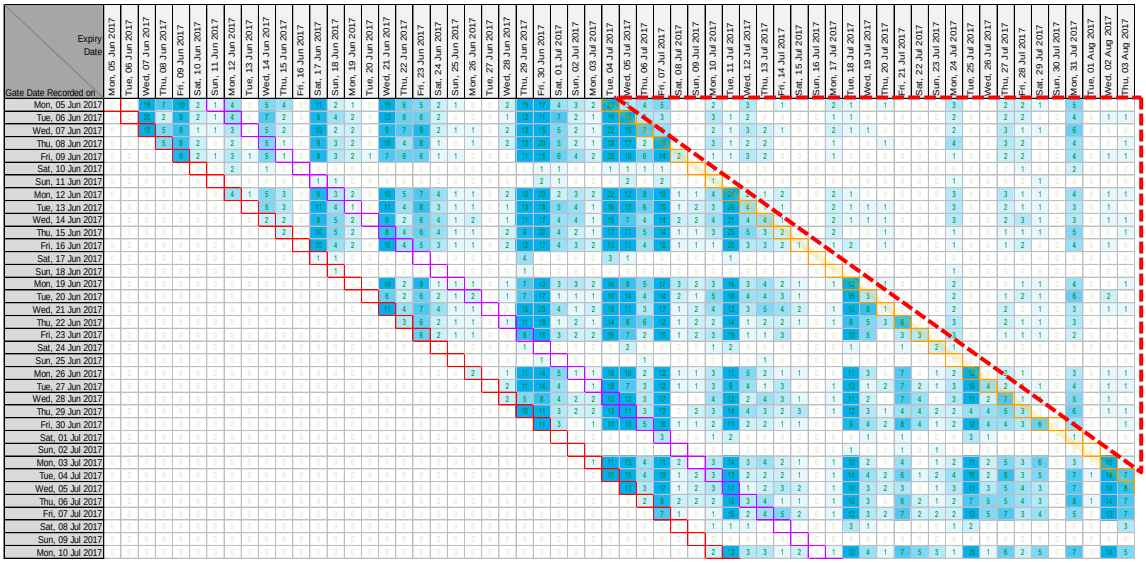
Annual Journeys per Ticket

5.21 The gateline assigns the same code to both Monthly and Annual tickets ("MT", with replacement tickets marked "ZT"). At this 'macro-processing' stage the outputs from the program are left in this form, Monthly and Annual Season tickets being left aggregated. When looking at the observations vs expiry date matrix for Monthly and Annual tickets in Figure 5-4, the Annual tickets can be seen in the upper right-hand side of the matrix where expiry dates are more than one month after the observation day.

5.22 The data in the red triangle are expected to be Annual tickets. However, this data is not exclusively Annual data as custom Monthly Season tickets can be purchased that have a validity of over one month.

5.23 To separate the data in a robust manner, ensuring we are only viewing Annual tickets, the Monthly ticket LENNON data is used. By excluding any observation days (rows on the matrix) where a Monthly ticket is valid, we can create a mask over the Monthly and Annual gateline data that excludes every observation which might have been Monthly Season ticket as determined from LENNON.

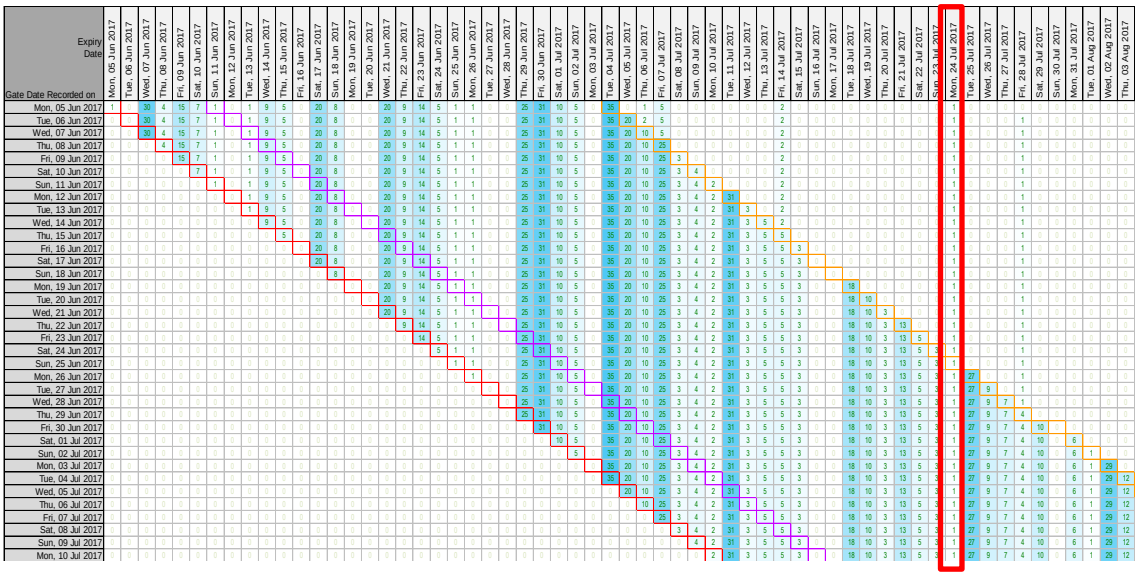
Figure 5-4: Monthly and Annual Ticket Observations



5.24 The LENNON data can be represented in the same observation vs expiry date matrix to allow comparison between the number of valid tickets and the number of observed tickets. In Figure 5-5, Monthly LENNON data is shown.

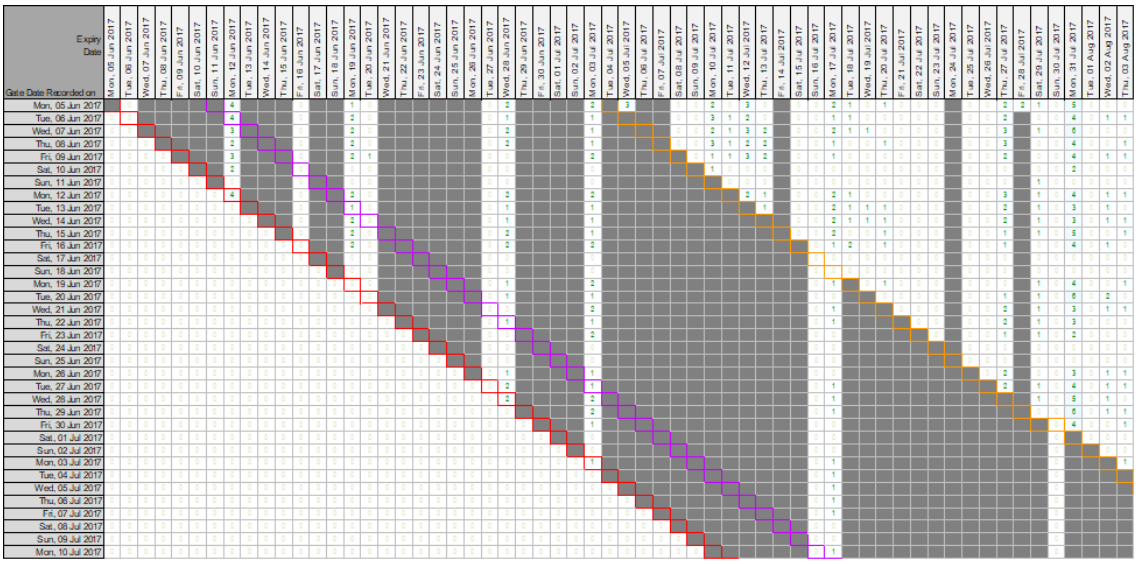
5.25 Note that there are some columns that have numbers extending into the upper right-hand section of the matrix, with one such example highlighted in the figure. Columns that have numbers extending into the upper-right hand side of the matrix are displaying where custom "Monthly" tickets are in existence.

Figure 5-5: Monthly LENNON Data Matrix



5.26 Using this LENNON data, we have created a mask to place on the Annual and Monthly gateline data matrix, where cells are covered if a valid Monthly ticket is in existence. As a result, we have covered all observations that could be a Monthly or an Annual, and we have been left with exclusively Annual observations.

Figure 5-6: Mask on Observations to leave exclusively Annual Ticket observations

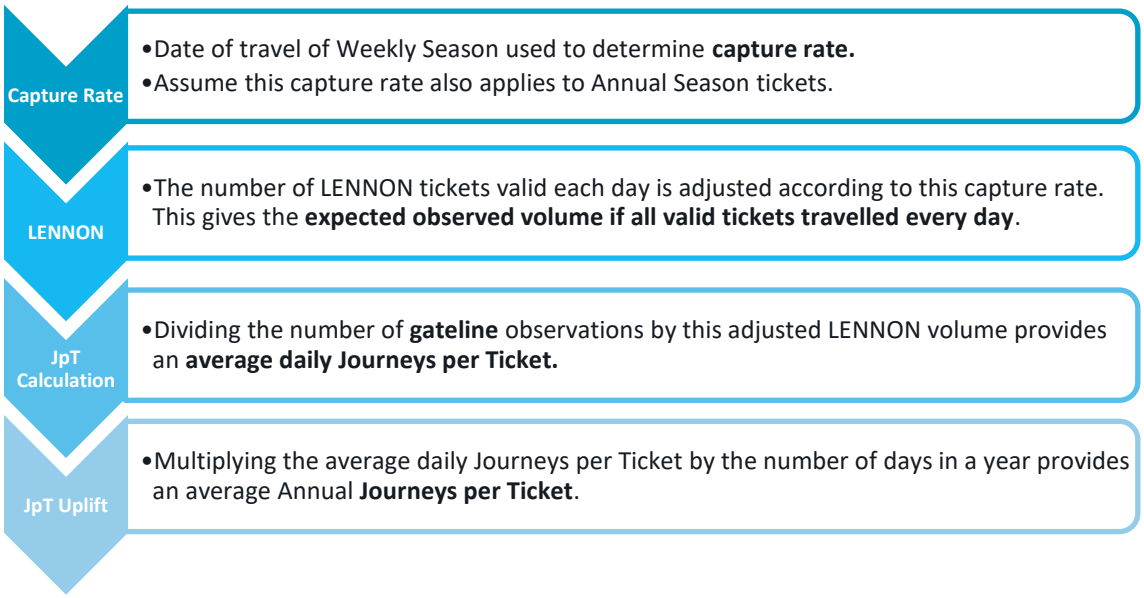


5.27 After finding the number of the Annuals observed on each day, we have calculated the Journeys per Ticket figure. Figure 5-7 displays the summary of the calculation approach, which is then discussed in more detail below.

5.28 Note that only a small part of a much large matrix is shown in the graphics above. The full matrix is 365 in height, and 730 (two years) in width, thus the majority of the data is not 'masked' by this process.

5.29 This gives us a clean observation of tickets which we know to be annual Season tickets, and we can deduce an equivalent dataset from LENNON with the number of ticket which were in circulation.

Figure 5-7: Summary of approach to calculating Annual Journeys per Ticket



5.30 The capture rate for a given day has been calculated using the Weekly gateline data, as described earlier in this Section. It has been assumed that the capture ratio calculated from the Weekly gateline data will represent the capture rate of Annual tickets.

5.31 Using the capture ratio that has been calculated for each expiry date, we have applied this rate to the number of LENNON tickets that were valid for each given expiry date. The capture rate adjusts the number of LENNON tickets which has provided us with the number of tickets that are expected to be observed, assuming all valid tickets travel every day.

5.32 To calculate the Journeys per Ticket on Annual tickets, we have first calculated the average daily journeys taken per ticket. This has been done by dividing the observed journeys by the capture rate adjusted LENNON data. Taking this average and multiplying by the number of days in the year has provided us with the average Journeys per Ticket for Annual Season tickets.

Refunded Tickets

5.33 As Annual Season tickets are closer to the end of their validity date, passengers may request a refund of the remaining portion of their Annual ticket. We have therefore adjusted our Annual JpT estimates to take into account this impact, by uplifting the initial estimates by the proportion of journeys allocated to refunded tickets relative to the total number of journeys on Annual tickets (both First and Standard Class), based on LENNON data.

5.34 This uplift has been calculated to be 4.8% of Annual Season ticket, consistent for both 2017 and 2018 and has been applied to the original Annual JpT estimates.

Replacement Tickets

5.35 After investigating the number of tickets that were replaced by reissued tickets it became apparent that, on average, passengers typically wait several before renewing their unusable Season ticket. This analysis could be done as reissued tickets are given a different ticket code by the gateline ("ZT" instead of the original "MT").

5.36 By investigating the gateline data observation matrices for exclusively "MT" tickets and "ZT" tickets an average delay time could be found between the loss of "MT" tickets from one matrix and the increase of "ZT" tickets in the other matrix. This was done by picking expiry date columns

that only contained observations of a maximum of one "MT" ticket and one "ZT" ticket. By then searching for the last observation of use of the "MT" ticket and the first observation of "ZT" use we can calculate the number of days in-between.

5.37 Using particular flows, which had been found to have a consistently high capture rate, and by comparing the first issue observations with the reissued ticket observations, we found that an average of 12 days passed between the last use of a first issued ticket and the first use of a reissued ticket.

5.38 Our investigations also suggested that approximately 40% of Annual tickets are reissued, or fail, over the course of the year. Note that this figure has been derived based on a relatively small sample of gateline data, and then using only dates where a single annual ticket was sold for a flow. This figure should be treated with caution, but given the relatively small impact from its use here it has it is considered to be sufficient.

5.39 The failure rate of Annual tickets was calculated by taking the ratio of the number of "ZT" tickets to the number "MT" tickets observed in the Annual data. This ratio was calculated for every expiry date where we had the largest number of observation days for valid Annual tickets.

5.40 This means that for two out of every five tickets (representing the proportion of reissued tickets), we expect that 12 days of observations are lost due to the passenger not having a ticket that will work at the gate line. This equates to approximately four and a half days of observations are lost per Annual ticket, i.e. 40% of 12 days of observations.

5.41 As a result, we expect the number of annual journeys per ticket which we originally calculate to be approximately 1.2% lower (4.5 days over 365 days) than in reality due to travel on unreadable tickets (which is therefore not visible to the gateline) before a passenger obtains a replacement. We have applied this percentage as an uplift to our annual journeys per ticket results.

Monthly Journeys per Ticket

5.42 To exclude the Annual Season ticket observations from the gateline data, from the combined Monthly and Annual ticket observations, it was not possible to use the same 'masking' approach that was used for the Annual tickets. This is because in the final month when monthly tickets are valid annual tickets share the same expiry date are always also valid.

5.43 To calculate the Monthly Journeys per Ticket, the approach instead was to forecast the usage of Annual tickets in their last month of validity, when we would expect to observe a mix of Monthly and Annual Season ticket usage.

5.44 To do this we calculated a normalised 'shape' which shows the typical pattern of how many Annual tickets were observed each day. This was calculated by normalising the number of observations of Annual tickets on every observation day. We call this shape the 'Normalised Busyness Index'. We can scale this shape for the number of annual tickets observed travelling, prior to the sale of monthly tickets, for any given expiry date.

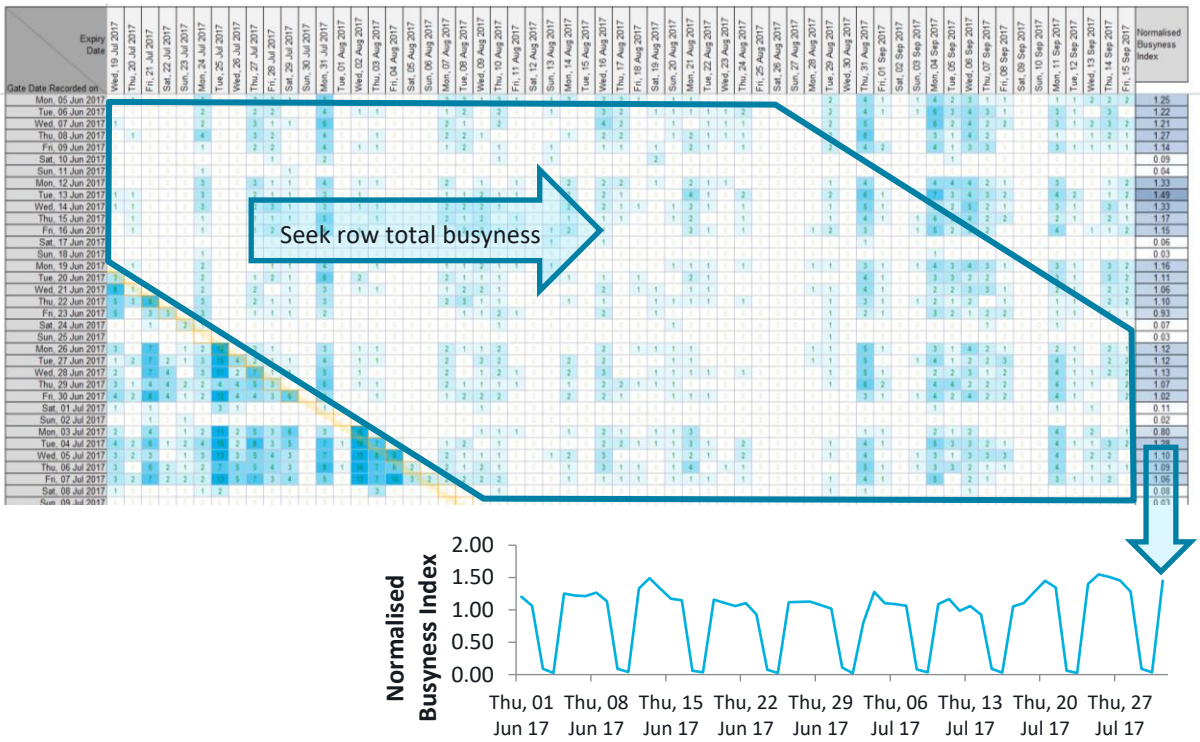
5.45 To find the busyness shape, the total volume observed for all expiry dates more than a month after the observation date is found for each day (i.e. excluding normal monthly tickets and seeking instead to look at how many annual ticket holders are travelling on each day). Since we are not interested in the absolute number, just the relatively popularity of travel on each day, we then scale this profile so the average is one trip per day.

5.46 A full year profile is generated, and this profile will typically show more travel occurs on weekdays, while travel is usually lower on weekends and bank holidays. Seasonal effects can

also be observed within this index. This shape can then scaled or "normalised" to match typical annual season ticket usage for any given expiry date; furthermore, and since a whole year is known, it can be used to inform analysis of the period of time when there is a mixture of annual and monthly season tickets travelling. The process for generating this shape is presented in Figure 5-8 below.

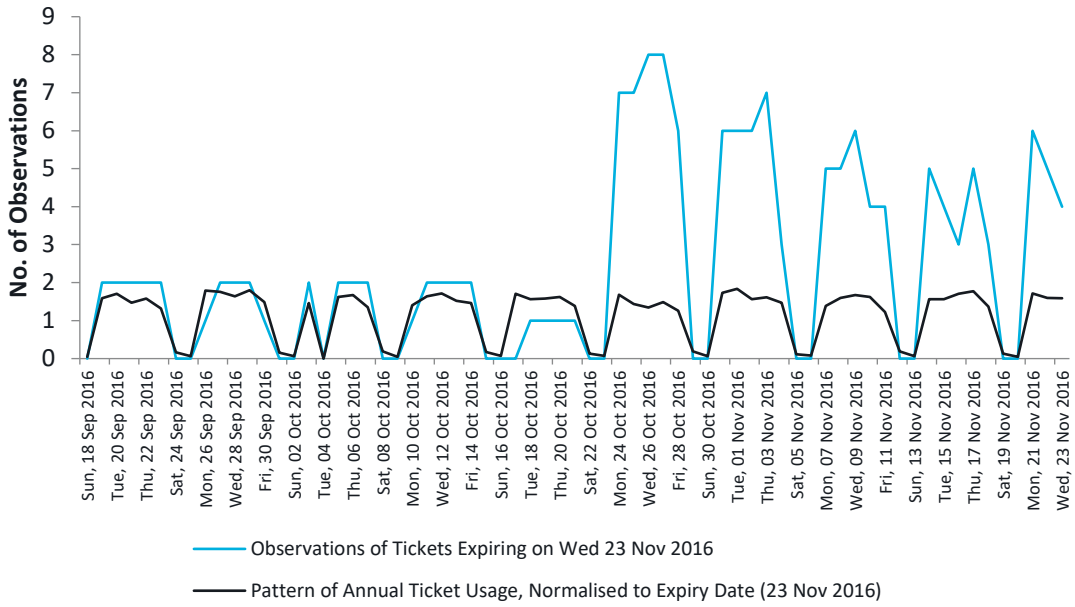
5.47 The difference between this scaled Normalised Busyness Index and observations for both Monthly and Annual tickets will be assumed to be only Monthly tickets. In Figure 5-8 we show an example of this shape that is produced for each observation day.

Figure 5-8: Normalised Annual 'Busyness'



5.48 By then picking a specific expiry date we can then apply the shape to the average number of Annual ticket observations that have this expiry date, to create an adjusted shape of Annual ticket usage for each expiry date. As the Annual Busyness Index includes every day we can project the adjusted shape of Annuals into the last month of validity, i.e. when we expect to begin observing monthlies. Any increase in observations above the Annual ticket observations we conclude must be Monthly Season tickets.

Figure 5-9: Annual busyness shape and Monthly Ticket observations



5.49 The number of Monthly observations has been calculated by subtracting the Annual busyness from the total number of observations, for each observation and expiry date. Once the number of Monthly observations has been obtained, it was possible to use the same approach described in the calculation of Annual tickets.

5.50 The capture ratio, calculated in the Weekly Season tickets Section, has been applied to the Monthly LENNON data to factor the LENNON data down to the expected number of Monthly observations if every Monthly ticket was used every day. By dividing the number of observations by the revised LENNON data we arrived at the Journeys per Ticket for Monthly tickets.

6 Stage 4 – Validation of Gatelines and Flows

Criteria for Validation

6.1 We have undertaken the analysis on almost twice the number of flows (360) than the originally proposed quantity (200), a number deemed a robust sample size for this study. This meant we could exclude anomalous results whilst retaining sufficient flows to maintain this robustness.

6.2 There are **four** criteria for the exclusion of flows. Table 6-1 describes these reasons and Table 6-2 shows the number of flows excluded for each reason for each Season ticket validity period.

Table 6-1: Description of flow exclusion criteria

	Exclusion Reason	Description
1	Low Number of Issues	The calculation of the Journeys per Ticket estimate for Weekly tickets is reliant on having a significant number of issues to give confidence in the calculated rate. The calculation of the Monthly and Annual Journeys per Ticket uses the capture rate of the Weekly tickets, so any error introduced in the calculation of the Weekly is compounded thereafter. Therefore, we have excluded any Monthly flows with fewer than 25 issues and Annual flows with fewer than 10 issues.
2	Low Capture Rate	The calculation of the Journeys per Ticket estimates is reliant on having a significant capture rate to give confidence in the calculated rate. Again here, we have excluded flows which we feel the gateline is not capturing the number of tickets which LENNON knows are in circulation from ticket sales. We have therefore excluded flows with a capture rate lower than 20%.
3	PTEs and Smartcard Areas	Areas within Public Transport Executives (PTEs) such as the London Travelcard Area in London due to the significant use of Oyster and contactless payment cards which the same level of information is not available through the gateline compared to paper tickets. It is a similar case for Huddersfield and the Yorkshire PTE. A selection Network West Midlands flows have been retained (e.g. into Coventry) as it was assumed that Swift uptake is low and does not distort the sample.
4	Grouped Stations	Flows which contain tickets where multiple destinations are valid were not excluded but grouped together, bringing the total number of flows down. We have therefore grouped all the BR flows to account for this.

Table 6-2: Summary of number of flows excluded by reason

	Exclusion Reason	Weekly	Monthly	Annual
1	Low Number of Issues	0	27	100
2	Low Capture Rate	22	22	22
3	PTEs and Smartcard Areas	0	11	11
4	Grouped Stations	23	23	23
	Total number excluded (and number of flows remaining)	45 (315)	83 (277)	156 (204)

6.3 Table 6-2 above shows that for all different periods of validity, we have retained more than the originally proposed 200 flows which were deemed a robust sample size for this study.

High-Level Validation

6.4 We have validated the results against the PDFH categories using RUDD, shown in Table 6-3.

Table 6-3: Comparison of selected flows compared to RUDD Season ticket totals for each PDFH flow type category

Code	Category	Total RUDD Journeys	Total In-scope Journeys	% Included
1	London South East	243,519,812	14,048,305	5.77%
2	London Rest of the Country	7,214,997	544,218	7.54%
3	South East Inter Urban	30,072,595	5,429,944	18.06%
4	Non-South-East Inter-Urban	25,409,949	4,546,532	17.89%
5	South East Non-London	1,221,209	183,519	15.03%
6	Non-London	28,304,588	2,507,987	8.86%

6.5 We have validated the results against the distance bands using RUDD, shown in Table 6-4

Table 6-4: Comparison of selected flows compared to RUDD totals for each journey distance band

Code	Category	Total RUDD Journeys	Total In-scope Journeys	% Included
1	<20 Miles	200,929,348	11,104,799	5.53%
2	20-100 Miles	132,849,555	16,077,845	12.10%
3	>100 Miles	1,964,247	77,860	3.96%

6.6 The proportion of observations for each of the categories in-scope, relative to the total RUDD journeys, has a good representation of the total Season ticket flows in GB, with no flows having a particularly low or high degree of representation, which indicates that our weighted average results will be significant and representative.

7 Results and Key Findings

7.1 We have generated Journeys per Ticket figures for the three key dimensions outlined in the methodology above: by period of validity (Weekly, Monthly, Annual), distance band (<20 miles, 20-100 miles and >100 miles) and by PDFH geography.

7.2 We have also generated a high-level figure for Weekly, Monthly and Annual for comparison purposes to the existing journey factors in LENNON. All of the findings support the hypothesis that Season ticket usage has reduced since the development of the existing factors in LENNON, likely due to changes in both remote and flexible-working behaviours impacting the needs and timings of commuting journeys.

Key Findings

7.3 The high-level results by period of validity are displayed in Table 8-1. As expected these are below the existing Journeys per Ticket estimates in LENNON.

Table 7-1: High level Journeys per Ticket figures

	Weekly	Monthly	Annual
New Journeys per Ticket	8.5	37	378
Existing LENNON factors	10.3	45	480
% reduction	18%	19%	21%

7.4 The following tables show the Weekly, Monthly and Annual (respectively) Journeys per Ticket results disaggregated to the dimensions outlined above. As expected, Journeys per Ticket are higher for commutes which involve urban environments (particularly London) and shorter distances. For long distance flows, we see a Journeys per Ticket figure of around less than half of the existing Journeys per Ticket estimates in LENNON.

Table 7-2: Weekly Journeys per Ticket results

	<20 Miles	20-100 Miles	>100 Miles
London South East	8.8	8.3	[12]*8.3
London Rest of the Country	*6.8	6.8	5.1
South East Inter Urban	8.8	8.2	*8.2
Non South East Inter Urban	8.7	7.9	5.2
South East Non-London	8.7	8.2	*8.2
Non-London	8.3	8.2	*8.2

Table 7-3: Monthly Journeys per Ticket results[13]

	<20 Miles	20-100 Miles	>100 Miles
London South East	42	32	*32
London Rest of the Country	*29	29	22
South East Inter Urban	38	38	38
Non South East Inter Urban	33	33	33
South East Non-London	37	37	37
Non-London	39	39	39

Table 7-4: Annual Journeys per Ticket results

	<20 Miles	20-100 Miles	>100 Miles
London South East	449	306	*306
London Rest of the Country	*271	271	197
South East Inter Urban	408	408	408
Non South East Inter Urban	325	325	325
South East Non-London	351	351	351
Non-London	338	338	338

7.5 For the non-London JpT estimates, we do not propose different values for Monthly and Annual Season tickets based on the defined distance bands (the same value is displayed in the above table for all distances). This is because for non-London flows, there is not a clear correlation to distance. This correlation is shown more clearly for Weekly tickets and we have therefore proposed disaggregated values for non-London tickets.

7.6 One possible reason is that the passenger decision-making for Monthly and Annual tickets is more difficult to consider given that the information on the likely frequency of travel for the

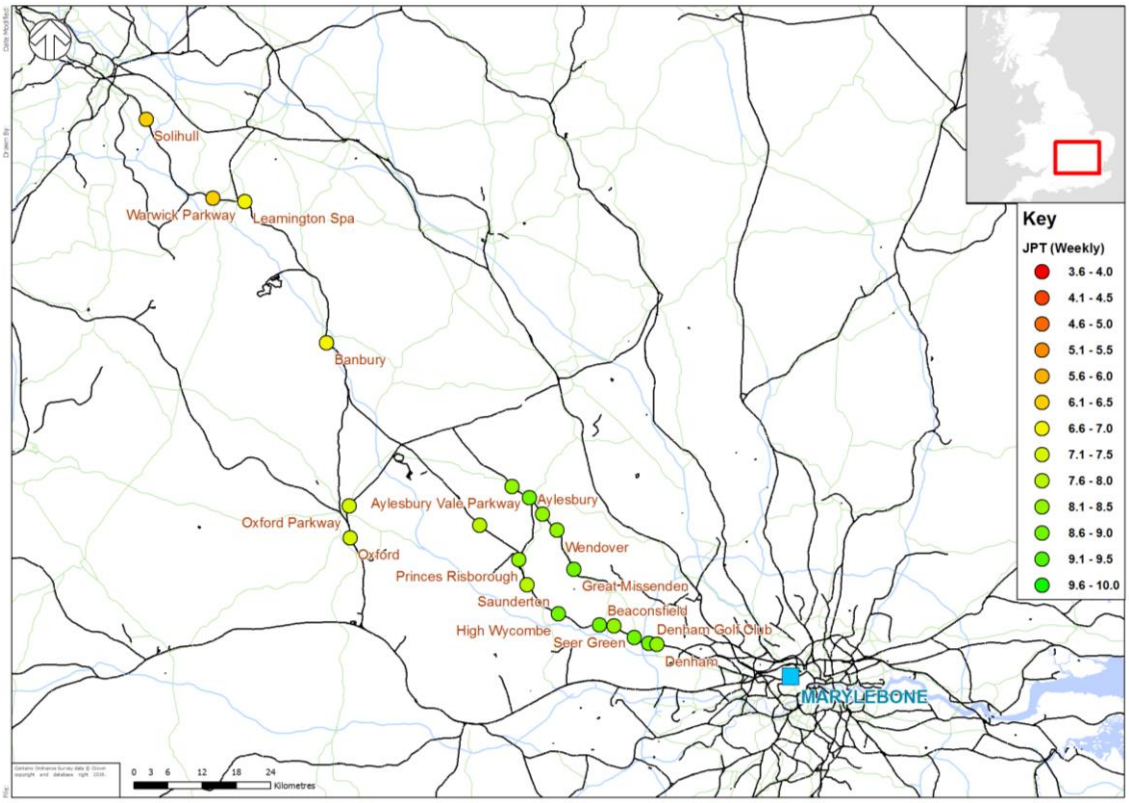
[12] For the categories with a low sample size, we have not estimated a JpT figure based on the flows within that category. For completeness, we have used on those cases the JpT values of the closer distance band for a given flow type. These are shown in italics and preceded by an asterisk. Generally, where the value has been taken from a longer distance band, this will represent an underestimate and, reversely, where it has been taken from a shorter distance band, this will represent an overestimate. However, it is not likely that a significant amount of flows will be found under each of these categories.

[13] It should be noted that these recommendations apply only to Monthly tickets. For tickets with a validity greater than a month but lower than a year, multiples of the values provided in this table may be used.

entire period at the beginning of this will be less complete and the trade-off with other products is more difficult to assess.

7.7 The results can also be displayed graphically on a map to further demonstrate the distance relationship. While we have generated these for each station in the analysis, Figure 7-1 shows the results for London Marylebone Weekly Season tickets. This confirms how the observed journeys taken on a Season ticket drop as distance of commute increases. Appendix B contains maps for all 15 of the selected stations.

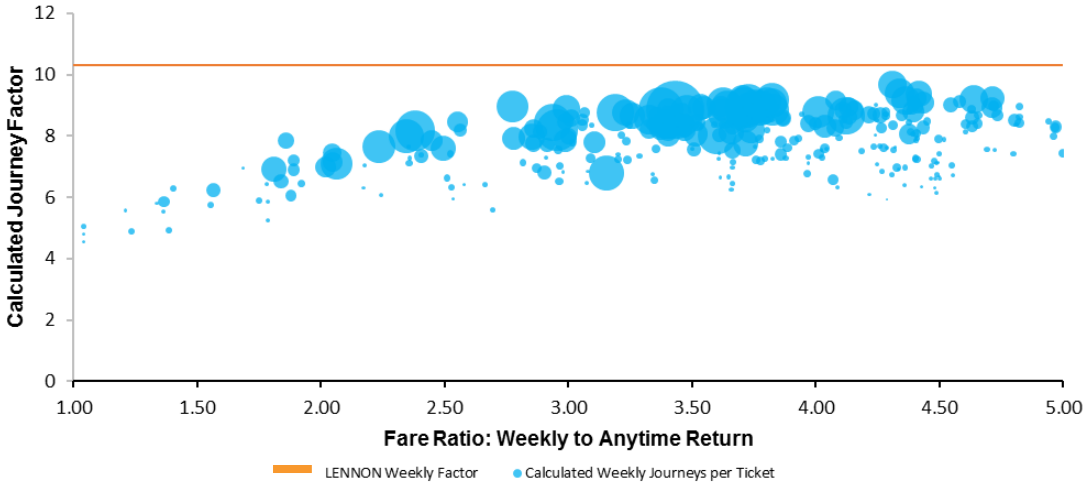
Figure 7-1: Map of Weekly Journeys per Ticket by line of route from London Marylebone Station



7.8 We have also plotted the full sample to demonstrate the strength of correlation between the Journeys per Ticket for each period of validity to the fare ratio.

7.9 Figure 7-2 shows the intuitive relationship between the two; for a higher ratio of Weekly Season ticket fare to the price of an Anytime return, we observe more journeys taken on the Season ticket, and for a lower fare ratio we see fewer journeys made, with fewer journeys required to make the Season ticket value for money. It is key to note that the existing journey factors in LENNON are higher than all flows in the analysis, and this does not differentiate by the key determinants set out here.

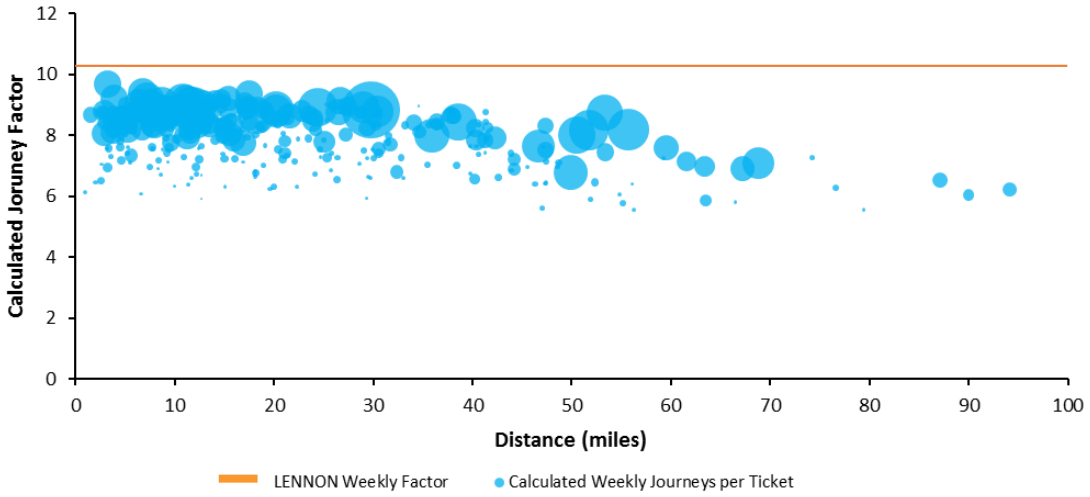
Figure 7-2: Correlation between Weekly Journeys per Ticket and Fare Ratio (Anytime Return and Weekly Season)



7.10 The Weekly JpT figure increases in line with the fare ratio until a value of around 3 for the ratio is achieved, after which the figure is relatively stable, indicating a cap on the average level of commuter travel irrespective of the fare ratio.

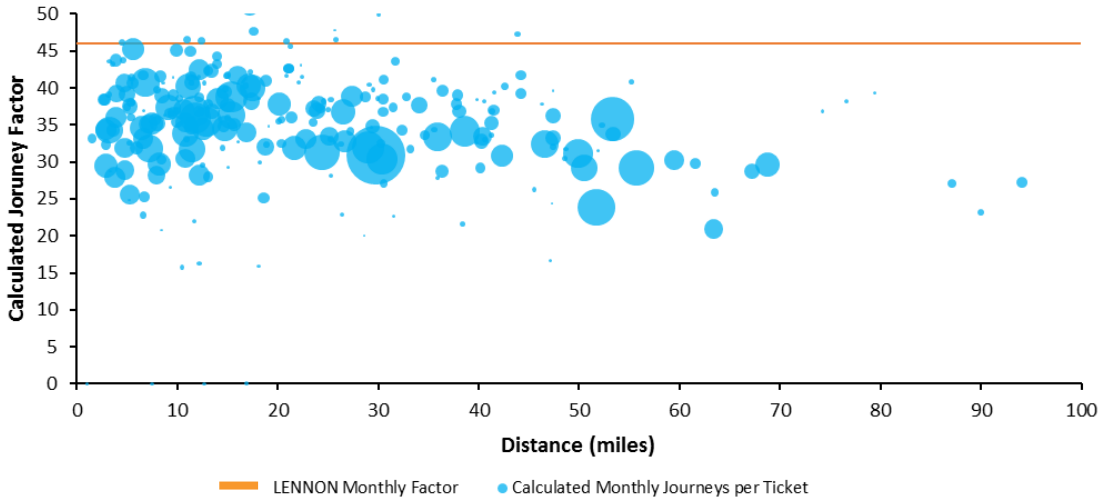
7.11 We have developed similar charts to demonstrate the relationship between journeys per ticket and distance. Figure 7-3 shows the relationship for Weekly Seasons which is quite evident, with almost all observations being between 6 and 10 Journeys per Ticket. It is again key to note here that all values estimated through our method are lower than the currently assumed LENNON journey factor.

Figure 7-3: Correlation between Weekly Journeys per Ticket and distance of journey



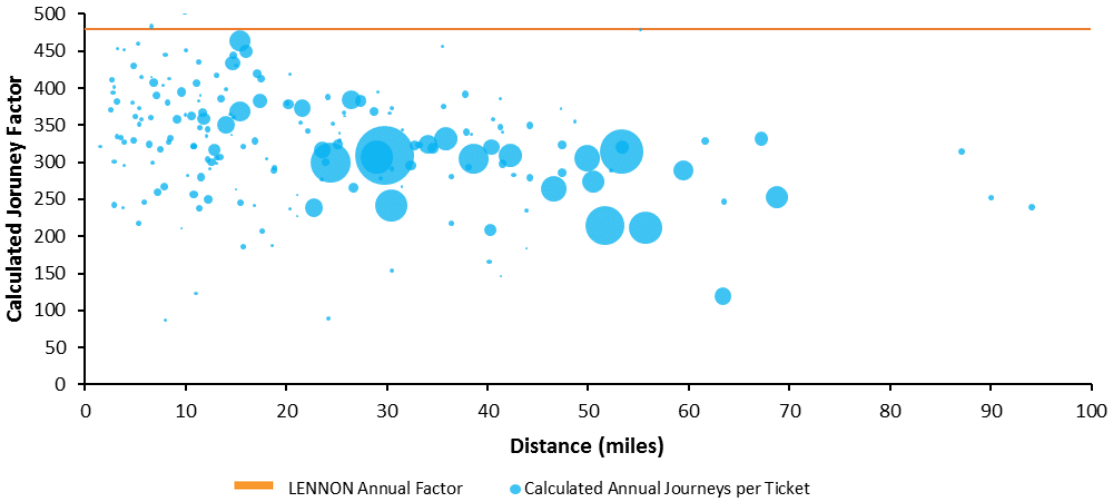
7.12 The relationship for Monthly Seasons shows greater variation with the observations more spread, indicated in Figure 7-4, yet showing similar correlation with distance.

Figure 7-4: Correlation between Monthly Journeys per Ticket and distance of journey



7.13 Again, the Annual Journeys per Ticket correlation to distance is not as strong as for Weekly tickets, shown in Figure 7-5, although it nevertheless displays the trend of fewer journeys over longer distance commutes.

Figure 7-5: Correlation between Annual Journeys per Ticket and distance of journey



Confidence Intervals

7.14 Confidence intervals (CI) were calculated for the Weekly, Monthly and Annual JpT nation-wide estimates at the 95% level of confidence. These demonstrate how confident we can be that the 'population mean JpT' (i.e. the average JpT value for all flows, not just those in consideration in this study) lies between the lower bound and the upper bound of the CI. Some flows may have JpT estimates that lie outside the calculated intervals – however, the mean across all flows has a 95% chance of falling within the range.

7.15 As described in earlier Chapters of this report, the process to select the flows used for this study consisted of picking 15 stations around country that fall under Network Rail categories A or B, which would ideally allow us to capture various flows for different distances and geographical categories. The top flows for each station, based on Seasons ticket issues, were then chosen to

form the sample for this study. When calculating the CI, we had to ensure that each flow was appropriately weighted by two factors:

- The number of journeys (attributed to the given ticket type) according to LENNON data; and
- The representation of the flow's geographical category and distance band (e.g. London and South East, greater than 200 miles) according to RUDD.

7.16 To calculate the CI, there is also an implicit assumption that the flows included in the sample cover a range of flows and do not have any inherent biases that might affect the JpT results for each of the ticket types.

7.17 To obtain the desired CI, it was necessary to compute an unbiased weighted estimator of the sample variance. This was done using the formula displayed below, where w_i is the weighting value for the specific flow, x_i is the flow's JpT estimate and $\bar{x}^*$ is the weighted average JpT of the sample.

$$s^2 = \frac{\sum_{i=1}^{N} w_i}{(\sum_{i=1}^{N} w_i)^2 - \sum_{i=1}^{N} {w_i}^2} \cdot \sum_{i=1}^{N} w_i (x_i - \bar{x}^*)^2$$

7.18 The weighting value (w_i) is calculated for each flow and ticket type (Weekly, Monthly and Annual) as a combination of:

- The number of journeys; and
- The proportion of flows in RUDD that fall into the flow's category of geography and distance.

7.19 The weighted average JpT is calculated using the following formula:

$$\bar{x}^* = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i}$$

7.20 Once the variance $s^2$is calculated, the standard error is derived and multiplied by the 't-stat value' for a normal distribution (which is 1.97 for the 95% confidence level) to give us the CI bound that we will both subtract and add to the JpT mean.

7.21 We have therefore applied the calculation described above to estimate the following confidence intervals for each of the nation-wide JpT estimates, as presented in Table 7.5 below:

Table 7.5: Confidence Intervals for JpT estimates, at a 95% level of confidence

Season Ticket Type	Mean JpT Estimate	95% Confidence Interval
Weekly	8.5	8.4 – 8.6
Monthly	37	36 – 38
Annual	378	365 – 391

7.22 The JpT mean value for the whole country is therefore 95% likely to be found within the above confidence intervals. It should be noted that this does not imply that there is a 95% chance that the JpT mean value for any individual flow or set of flows will be within that confidence interval.

8 Quality Assurance

Overview

8.1 Throughout the process we have checked and reviewed both the process and the emerging results from the analysis. This can be broadly categorised into four areas:

- The theoretical principles and process for calculating the Journeys per Ticket estimates;
- Spreadsheet Modelling Best Practice (SMBP) and the practical element to the analysis;
- The selection, categorisation, validation and review of stations and flows in the study; and
- A series of internal meetings and quality assurance reviews to comment and feedback on outputs.

Calculation of Journeys per Ticket estimates

8.2 The principles and analytical background for the processing of gateline data has its foundations in multiple existing studies using gateline data for assessments of demand, predominantly for Chiltern Railways at London Marylebone Station.

8.3 Most of the Visual Basic (VBA) code[14] underpinning the analytical framework are therefore not new, have undergone multiple iterations and have been vouched for by a number of project teams, subject to thorough Quality Assurance processes.

8.4 In addition, all workbooks submitted as part of this study have gone through our Quality Assurance processes.

8.5 Furthermore, these have been shared with the Department for Transport during the development of this study and we have addressed their comments where appropriate.

Bidirectional JpT estimates within same flow

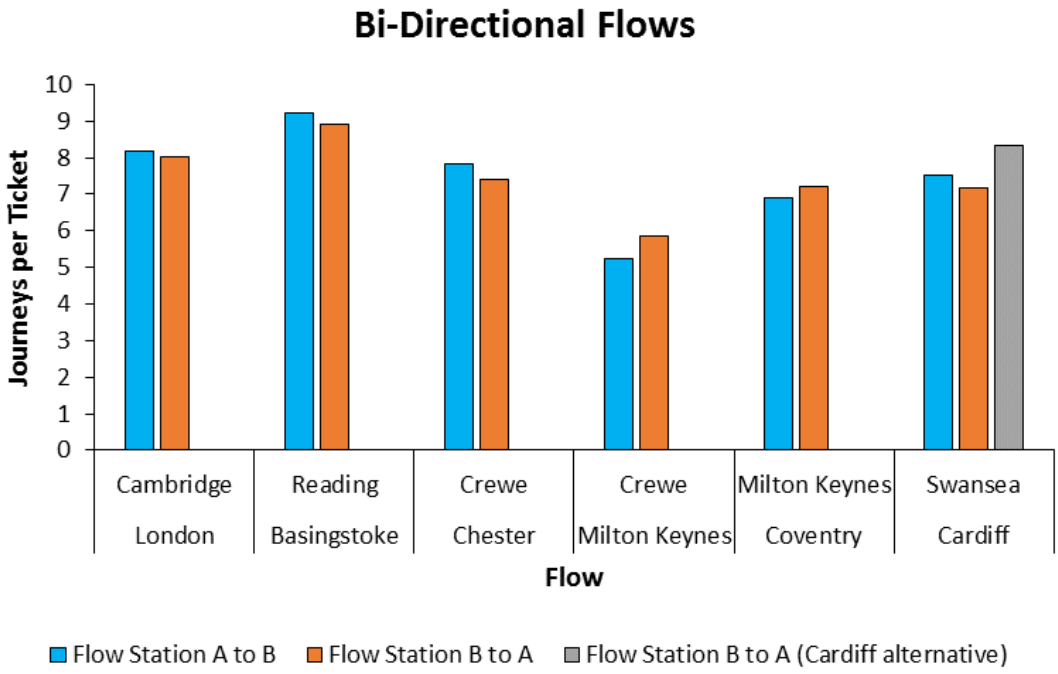
8.6 The combination of gatelines and flows selected means there are six flows for which we have calculated a JpT at both ends of the flow. This was possible for:

- Cambridge-London (Liverpool Street)
- Reading-Basingstoke
- Crewe-Chester
- Crewe-Milton Keynes
- Milton Keynes-Coventry
- Swansea-Cardiff

[14] The VBA code has been used to “mine” the gateline data and store it in Excel workbooks in a format which is practically easier to work with. This VBA code had previously been used for other purposes and had therefore gone through Quality Assurance processes.

8.7 This is useful to validate the process, given we would expect the average journeys per ticket to be the same in both directions (assuming all passengers use the same mode/route in both directions). The results for the Weekly bi-directional flows are shown in Figure 8-1 below.

Figure 8-1: Weekly JpT for Bi-Directional Flows



8.8 The calculated factors are similar in each direction, with a variance of between 2% (Cambridge-London) and 16% (Swansea-Cardiff). In the case of the latter, we have calculated 3 factors (covering both Cardiff Queen St and Cardiff Central).

Refunds and replacement tickets calculations

8.9 We have quality assured the calculations undertaken to estimate the uplift factors applied to Annual JpT factors corresponding to refunds and replacement tickets. These have been provided to the Department for Transport.

Spreadsheet Modelling Best Practice (SMBP)

8.10 The project team is highly experienced at delivering spreadsheet models which conform to SMBP principles, having prepared these for the Department for Transport on many occasions.

8.11 Our SMBP principles align with those of DfT's Analytical Assurance Framework and we have included this guidance in our processes and outputs.

8.12 In particular, the transposition of gateline data into workbooks has been undertaken automatically by a VBA macro, which has been subject to Quality Assurance. In addition, the transposition of data between the workbooks and the final report has been verified by the Project Manager once the tables and graphs have been generated, before the report has been issued as a formal output.

8.13 A series of auditing tools and cell mapping tools have been run on the workbooks used for the analysis, and the Results workbook submitted to the Department for this study is SMBP compliant and in line with DfT's Analytical Assurance Framework.

Flow Selection, Categorisation and Validation

8.14 Throughout the process we have undertaken reviews of the qualitative elements of the analysis involving the selection, categorisation, validation and review of stations and flows.

8.15 The Department's Rail Usage and Demand Drivers Dataset (RUDD) was used to validate and weight the results against total levels of demand, to make our results representative and validated against a dataset that has gone through in-depth Quality Assurance processes.

8.16 PDFC members were engaged early in the process with thoughts and comments on the project, and were requested to raise any concerns about the stations or flows which were selected, as outlined in Sections 3.7 to 3.11.

8.17 The project team have used the emerging results of the study to exclude flows which display inconsistent behaviour, show low capture rates, are grouped with other stations or other anomalies which would impact the sample as outlined in Section 7.

Internal Meetings and Reviews

8.18 The project team has met frequently formally, and there has been constant informal discussion and review throughout the development process to discuss both the methodology and the findings. In addition, the project team has engaged regularly with the Department for Transport, sharing early results and addressing their comments early in the process to avoid iterations at later stages.

8.19 In particular, as part of our Quality Assurance process, all formal outputs of this study, including this report, have been thoroughly reviewed and approved by the Project Manager and Project Director, ensuring all quality standards have been achieved.

Quality Assurance Actions

8.20 We have summarised the Quality Assurance undertaken on the analysis workbooks in Table 8.1 below. This includes who has developed originally each workbook, who has checked it and who has independently reviewed it, where appropriate. It also shows whether each workbook has been run through our cell mapper tool to verify compliance with SMBP principles.

8.21 All the spreadsheets named below have been provided alongside this report to the Department for Transport.

Table 8.1: Summary of Quality Assurance Actions

Workbook	Cell Mapped	Developer(s)	Internal Checker	Reviewer	Notes
NBK Fast CSV Import Filter 18 - All Flows.xlsm	No	NK	-	MB	Based on an equivalent workbook developed by MB this is an extremely optimised workbook, developed to give the same results but fast enough to it to be included in the iteration loop performed by the "Single Flow Analysis" workbook.
JPT DataMiner_v2.65.xlsb	Yes	OH	NK	-	Based on a (NK) workbook developed to perform a similar task for Chiltern Railways. Code and workbook structure was adapted for this task by OH with support and internal review by NK. Outputs have been sport checked vs the raw csv data.
Single Flow Analysis 69.xlsb	Yes	OH NK	NK	MB	Aspects of this calculation were initially developed in the JPT DataMiner workbooks -- however no calculations in these workbooks are now used. This workbook book was developed to allow all 360 flows to be modelled via the same calculation.
Journeys Per Ticket Results Workbook_v1.00.xlsx	Yes	NK	JP	MB	Superseding an earlier workbook "Journeys per Ticket_Emerging Results" this workbook consolidates and presents the results.
Journeys per Ticket_Refunds Impact_v1.01.xlsx	No	MB	AB	-	
ReIssued Ticket Comparison v5.xlsb	No	OH NK	NK		

9 Limitations and Caveats to our Analysis

Risks and Mitigations

9.1 Table 9.1 presents a summary of the main risks and mitigations identified in our analysis. The remaining of the section (from paragraph 9.8 onwards) describes in further detail the extent of these issues and their mitigations.

Table 9.1: Summary of main risks and mitigations in our analysis

Limitations	Mitigation	Comment	Post-mitigation Impact	Direction of bias
Non-recurrence of passengers' behaviour	Comparison of flow with and without competition	There is no apparent change in behaviour between both types of flows	Low	Not biased
Insufficient representativeness of flows	The sample sizes for each flow type are proportional to the market sizes	There might still be an issue linked to small absolute sample size, reducing the robustness of the estimate	Medium	Not biased
Attractor and producer stations – different behaviour	Comparison of gatelines in entry and exit modes. Comparison of Bi-Directional flows	Six JpT Bi-Directional flows obtained in this study and compared	Low	No clear direction of bias
Break of journey or split ticketing not considered	We have a limited number of attractor stations, where this effect is more significant	In these flows, high capture rates indicate that this is not a significant issue	Low	Potentially higher number of legs per ticket, but likely that estimation at key attractor station is correct
Refunds on Annual tickets not considered	Estimate the proportion of refunded Season tickets and derive a factor to apply to JpT estimates	Not implemented in this research. It is unclear whether journeys associated with refunded season tickets are reversed out	Low	This has been accounted for. No obvious remaining direction of bias

Multi-destination tickets set at same prices	Analyse flows where there is potential for this issue (e.g. Cardiff flows)	No evidence that this issue impacts the scale of the JpT estimates	Low	Not biased
Distance flow categorisation combining LENNON and RUDD	Classification of flows in distance bands (under 20 miles, 20-100 miles, over 100 miles)	Using LENNON rather than RUDD for distances is not likely to move flows to other distance bands	Not material	Not biased
Physical degradation of tickets not considered	We have accounted for reissued Season tickets	The inclusion of reissued tickets led to a 2% uplift was applied to the original factor	Not material	This has been accounted for. No obvious remaining direction of bias

Recurrence of Passengers Travel Behaviour

9.2 We have included flows on competitive routes (e.g. Oxford-London) on the basis that a passenger who uses their Weekly ticket on the first day on one route will continue to use that route for the remainder of the week. If there are passengers who switch mid-week, they are equally likely to switch from route A to route B, as they are from route B to route A, and therefore it is unlikely to impact the calculated JpT. We have compared flows with competition to those without, and the calculated Journeys per Ticket estimates follow the same trend in relation to fare relativities, and therefore we see no reason to exclude these flows. The assumption about competitive routes not only applies to different routes, but also to origin and destination choices and, equally, we do not consider that these would have an impact on the estimated JpT figures.

9.3 There is a further risk that passengers' behaviour in the use of gates changes throughout the life of a season ticket. This may relate to the degradation of tickets (covered in 9.16), or other changes (e.g. station familiarisation) which influence the use of gates. Whilst this may be true of some new passengers, the majority of new season tickets are likely to be held by passengers who have travelled from the station before (either on a season ticket or other ticket type), and therefore we do not expect any significant changes in behaviour across the market through the life of the season ticket.

Representativeness of Flows

9.4 The process we have adopted in selecting the flows to be analysed is an example of cluster sampling. We have divided the market for Season ticket travel into strata which display varying characteristics which would influence frequency of use (journey distance, geography and price relativity), and have selected a range of flows within each, where possible. The exceptions to this are:

- London <> Rest of Country, <20miles
- London <> South East, >100 miles
- South East Inter-Urban, >100 miles
- South East non-London, >100 miles
- Non-London, >100 miles

9.5 The total Season ticket demand for these flow types is 0.1%[15] - therefore the omittance of flows within these groups for the analysis is unlikely to have an impact on the final results.

9.6 The London <> South East, <20 miles category is the largest in terms of RUDD journeys. Our analysis has used 0.7% of these journeys, which is the lowest proportion analysed (with the exception of the categories with no data, shown above). This is due to the large number of flows which fall into this category, and the need to gather a cross-section of all geographies using a limited number of gatelines.

9.7 The exclusion of Smart ticket data from this research (see 10.17), which has high prevalence in this category, was a further limiting factor on the proportion of journeys analysed. However, we have no reason to believe that season ticket usage varies between Magstripe and Smart tickets. It is also likely that the frequency of travel in this category is largely driven by the requirements of employment within London, and the selected flows should be representative of this employment despite the restricted spread of the rail geographies. It should be noted that this study does not provide recommendations for Transport for London products and, in this context, smart tickets refer to rail operators smart season tickets.

Attractor and Producer Stations

9.8 A number of flows have been isolated which, relative to the overall capture rate of the gateline, have a low capture rate and do not face competition from other operators. These flows typically have a large station as the other (i.e. station where gateline has not been analysed) end of the flow (an example of this is Crewe – Manchester). In these cases, the gateline is at the producer end of the flow, and therefore the majority of Season tickets, with a gate in exit mode, will be captured in the evening on the return leg of the commute. Typically, there is a larger spread of timing of the return leg of a commute, and it is therefore likely that a greater proportion of passengers pass through the analysed station at times when the gate is left open. For Crewe – Manchester, the capture rate is 56%, compared with an average for Crewe of 73% (based on the top 24 flows). Whilst the calculated Journeys per Ticket estimates do not indicate that these flows are outliers, it is possible that the lower capture rate increases the variability in the results. A recommendation for future work in this area would be to either:

- Capture the observations for both gate in entry and gate in exit mode for each day, and select the maximum for each day;
- Capture both gate in entry and gate in exit mode, but limit the observations based on time (for example before midday, assuming all outward commuting has been undertaken by this time, but no return commuting).

Break of Journey

9.9 Season tickets allow “break of journey” at any stations on the route valid between the stations referenced on the ticket. Therefore, tickets may be used to travel to other stations, which we are not able to pick up in the analysis. This is particularly true where the gateline we have analysed is at the “attractor” end of a flow – we would expect every trip to start at the “producer” end, but may be used to travel to an intermediate station. An example of this would be an Oxford Parkway – London commuter, who uses the ticket to travel to Bicester at the

[15] Calculated from RUDD 2016/17 data

weekend. Whilst this would lead to underestimating the number of journeys undertaken on a ticket, the analysis is still suitable for calculating the journeys undertaken between the two named stations, which is the purpose of the Journeys per Ticket estimates used in LENNON.

9.10 Similarly, a Season ticket may be purchased as part of a split-ticket transaction to circumvent break of journey restrictions. In these cases, the ticket would not be seen at the gateline at one end of the journey. We have looked at the Manchester-Milton Keynes flow as an example of this, but the high capture rate indicates that Season tickets are not currently being used to split ticket.

Refunds on Annual Tickets

9.11 LENNON data used for this research was aggregated into Season ticket groups for Weeklies, Monthly+ and Annual. Inclusion of refunded tickets within this dataset has not been considered as part of this research, therefore the number of tickets in circulation from which the JpT figures are estimated might be over-estimated.

9.12 This is unlikely to have any significant impact on the weekly JpT, but may impact longer-term season tickets for which refunds are more likely. A future refinement of this work could analyse the profile and proportion of refunds over the duration of the term of season tickets to re-estimate the number of season tickets in circulation over time.

9.13 We have been unable to confirm whether the journeys associated with season tickets are reversed out from LENNON in the case of a refunded ticket, or whether it is only the revenue which is refunded. If they are not reversed out, the methodology used for this research mirrors that used in the LENNON data.

Multi-option Tickets Set at Same Prices

9.14 Season tickets to Cardiff Central and Cardiff Queen St are set at the same price, but are not priced as a grouped station (i.e. Cardiff Stations). Capture rates for Cardiff Central tickets are particularly low from the Valley Lines. This may be due to passengers purchasing Cardiff Central tickets, with the intention of only travelling as far as Cardiff Queen St, but with no financial penalty for purchasing the longer-distance ticket. As with Oxford-London, there is no evidence to suggest the Journeys per Ticket figure calculated for these flows is not robust.

Flow Categorisation

9.15 Distances have been obtained from the RUDD, when the flow has been contained within the dataset. RUDD uses MOIRA distances, which are calculated using primary locations. In cases where one of the stations for a flow would be a secondary location in RUDD, the distances will not be accurate. However, for the purposes of cross-tabulation, it is unlikely that this process will result in flows being assigned to the incorrect category.

Physical Degradation of Tickets

9.16 Season tickets are, towards the end of their validity, more likely to suffer from degradation than at the beginning. Therefore, the gateline capture rate for each ticket may reduce over its life. This will lead to the estimate of Journeys per Ticket estimates, particularly for Annuals, which are lower than in reality.

9.17 However, we have addressed this by considering “ZT” tickets representing replacement tickets and have undertaken analysis to understand what the likely timeframe is for people to request a new replacement ticket, and have factored this into our Annual ticket calculation.

9.18 Passengers who use rail tickets to travel on the London Underground might find these degrading at a faster rate than those who do not; however, given that our analysis excludes flows which are in-boundary and minimise Underground travel, this is not an issue for our methodology.

10 Conclusions

10.1 The main output of our work has been weighted average Journeys per Ticket estimates for each of the three Season ticket types, resulting in the following values:

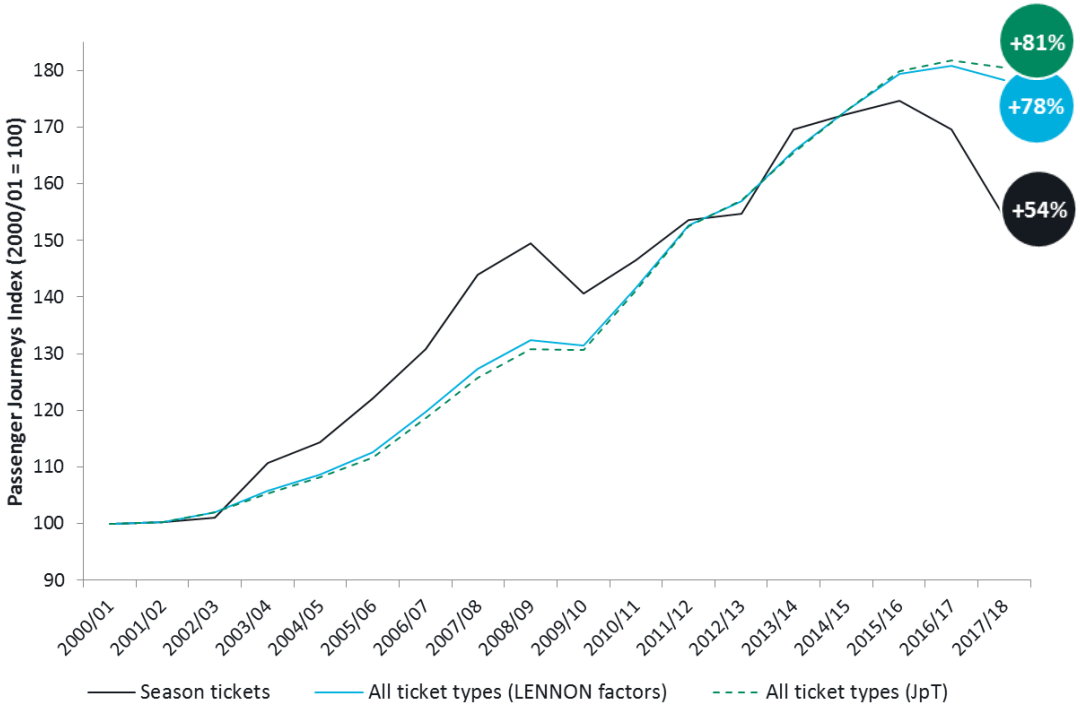
- A **Weekly** Journeys per Ticket figure of **8.5** (83% of the original 10.3 LENNON figure);
- A **Monthly** Journeys per Ticket figure of **37** (82% of the original 45 LENNON figure); and
- An **Annual** Journeys per Ticket figure of **378** (79% of the original 480 LENNON figure).

10.2 We have also estimated these JpT for each of the distance and geography categories we have defined, which shows that these are higher for London flows and decrease in line with the distance of the flow. In all cases, the inferred JpT are always lower than those assumed in LENNON, being significantly lower in long-distance flows.

10.3 As we have described in Section 1, in the recent years we have seen a decline in Season ticket journeys which, among other factors, has caused that overall rail journeys have also decreased. However, following the findings of this report, we have found that the number of journeys made on Season tickets is lower than what is assumed in LENNON, therefore the proportion of Season tickets is lower than currently assumed.

10.4 This results in the decrease in Season ticket usage having a lower influence in the overall ticket usage growth. We have replicated Figure 1-2 assuming that the journeys made on Season tickets is 80% of the current assumed number, which impact the overall journeys growth due to the decrease in the proportion of Season tickets (a 3% higher than currently assumed). This is shown in Figure 10-1.

Figure 10-1: Passenger journeys growth, with and without inferred Journeys per Ticket



Source: ORR Data Portal, Table 12.7 Passenger Journeys by Ticket Type and Steer/WSP analysis

10.5 The conclusions of our study confirm that the current Season ticket usage indicates a lower JpT figure than currently assumed in LENNON; this conclusion is consistent with trends outlined in other relevant industry studies. There are, however, some areas of opportunity for further work which might yield improved estimates in some geographies, for instance an assessment based on smart card data where penetration of these products is high.

Recommendations

10.6 We present below some potential recommendations inferred from this study, mostly relating to the potential implementation of its results, the gateline data as the main data source, and the research itself and its future improvements/refinements.

Implementation Recommendation 1

10.7 With the forthcoming changes to LENNON which will potentially allow for journey factor estimates to be specified by their users, our study represents a solid evidence base upon which these new factors could be based. However, the evidence underpinning the use of new factors may need to be reviewed on a case by case basis and may need to be tailored for application to specific purposes.

Implementation Recommendation 2

10.8 The Passenger Demand Forecasting Council (PDFC) may consider the creation of revised base matrices in MOIRA models to reflect these recommendations. The revised JpT estimates might impact the volumes upon which the MOIRA matrices are constructed, scaling the values for Season ticket journeys accordingly.

10.9 However, there might be some aspects of these MOIRA matrices that would need to be considered carefully, such as the infills for PTE areas, which might not be driven by the assumed

LENNON factors (e.g. might be derived directly from surveys to passengers). These would need to be considered should the matrices be considered for revision. In particular, it should be noted that the recommendations from this study do not apply to Transport for London season products.

10.10 When implementing these recommendations to Annual tickets, these need to be applied to the number of tickets prior to adjusting for refunded tickets. The impact on the JpT figure associated to refunds has already been adjusted for.

10.11 These recommendations do not apply to some infills in MOIRA representing tickets sold in Passenger Transport Executive (PTE) areas, where those infills were sourced directly from journey data.

Data Recommendation 1

10.12 The Rail Delivery Group (RDG) may desire to make gate data from Cubic and other sources more accessible for this type of research given the usefulness and wealth of gate data in addressing rail demand research queries.

Data Recommendation 2

10.13 In parallel, Cubic and other gate data companies may benefit from improving gate data systems, which record a more detailed set of product code data, which may be used for research purposes.

Data Recommendation 3

10.14 There is the possibility to develop a further understanding about the smart data currently recorded by different entities and explore its potential going forward, as well as provide recommendations about potential improvements to how it is recorded. These could be useful for future iterations of this research.

Research Recommendation 1

10.15 This piece of research should be replicated periodically to check whether/how the JpT estimates are changing over time. This would keep the inferred JpT information up to date and would also potentially support future changes in ticket purchasing behaviour.

10.16 This would allow to confirm its findings and to have a more robust set of estimates that may be implemented within the rail industry.

Research Recommendation 2

10.17 Uptake of Smart Tickets, especially for Season tickets, is likely to increase over the next few years. In the future, this may make the analysis of Smart Ticket data a robust alternative source of data when deducing Journeys per Ticket estimates.

10.18 However, it will also be hard to determine whether the estimates associated with passengers who use Smart Tickets are materially different from those who continue to use paper tickets, and this may make developing a robust understanding of journeys per ticket estimates on those flows where both options are available more complex.

10.19 In addition, increased prevalence of account-based ticketing would also mean that there is no longer the presence of a physical Season ticket, rendering the analysis unnecessary.

10.20 As smart ticket uptake increases over time, this research would benefit from being expanded to take into account flows where smart ticket usage is higher, so that we understand better the travel behaviour under this ticket types.

steer wsp

Appendices

steer wsp

Appendix A: Summary of PDFC Members Feedback

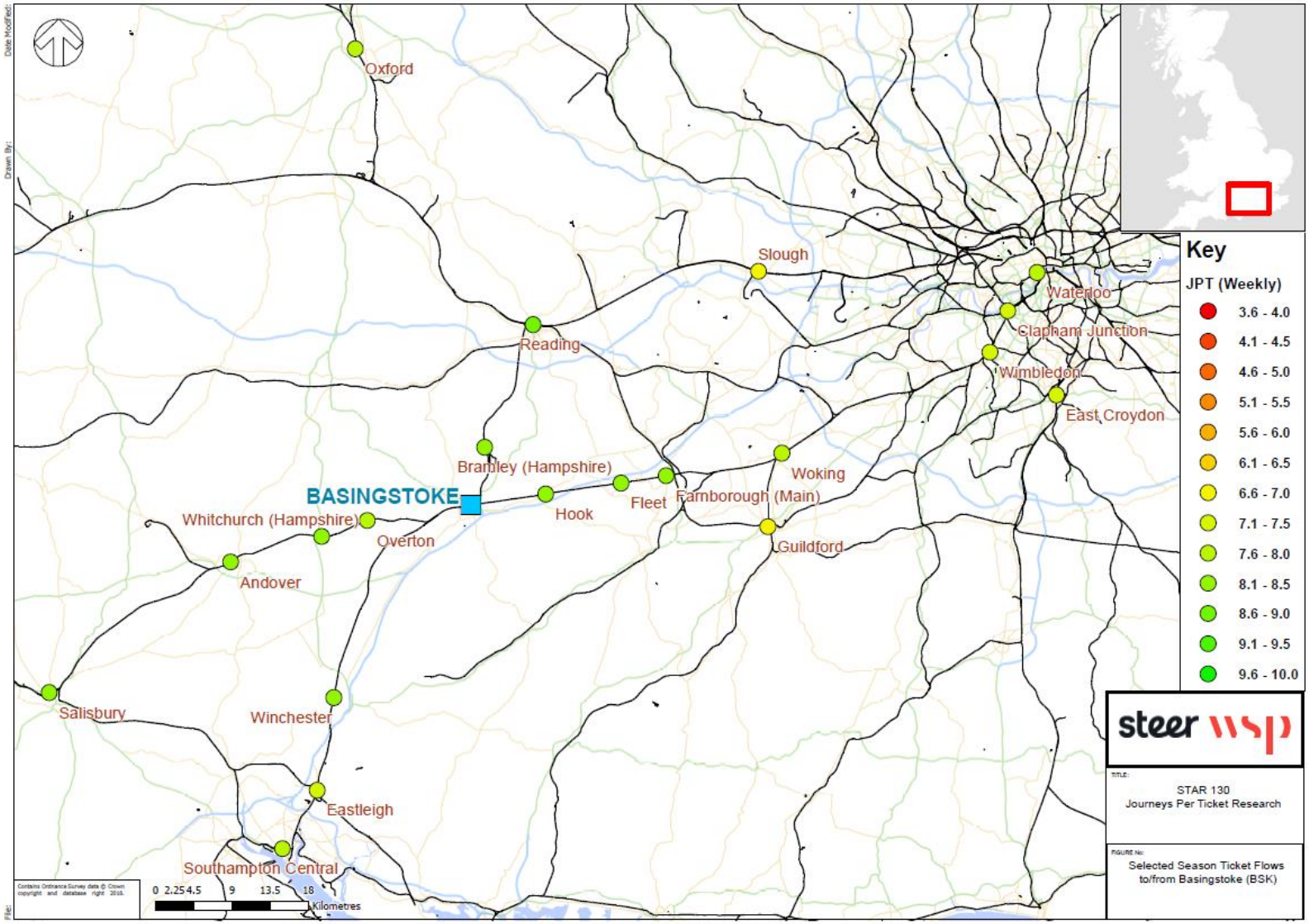
Author	Organisation	Date Received	Feedback	Our Response
Nigel Searle	Govia Thameslink Railway	27 March 2018	Include Gatwick Airport in the station selection and demand profile of different ticket types (Season, Full, Off-Peak).	Analysing demand profiles, while they drive capacity requirements on the network, is not part of the remit for this research. Similarly, this study focuses primarily on the behaviour of Season ticket holders. As such, we consider that passengers from/to **Gatwick Airport** do not have a particular commuter behaviour (i.e. there would not be many Season tickets holders), so this would not add any further robustness to our analysis.
Niki Hill	Govia Thameslink Railway	27 March 2018	Smart ticket data for Season tickets could be made available if considered useful.	Our methodology has been set out on the basis of using ticket gateline data. Smart ticket data, if available, would be a **useful source** and we would welcome it; however, **depending** on its **format** (and attributes that are included) we may or may not be able to incorporate it into our analysis as part of the current phase of the work.
Dean Hutchinson	South Eastern Railway	27 March 2018	London Terminals tickets (or Travelcards including Zone 1) will have a choice of London Terminal (Charing Cross, London Bridge, Waterloo East, Cannon Street, Victoria) with a few services to Blackfriars and High Speed to St Pancras. The preferred terminus cannot generally be inferred from the LENNON data so the validation will not be possible. Consideration could be given to St Pancras instead of Charing Cross.	We agree that selecting Charing Cross would increase the complexity (and uncertainty) when validating with LENNON due to the issue described. Therefore, we propose using **Liverpool Street to replace Charing Cross**. Although tickets to Liverpool Street would also be London Terminals tickets, there would only be a few interchange points, e.g. Stratford or Tottenham Hale. It would not be financially sensible for passengers to buy a London Terminals ticket and then interchange in Stratford as opposed to buying a Stratford ticket and then using the tube, therefore we consider Liverpool Street would capture most of the journeys on London Terminals tickets. Note: Our methodology does not require (or even anticipate) that all passengers on a LENNON flow will be observed at the chosen gatelines; however, we do need travel patterns to be consistent. Our approach considers that either we always see a passenger at the gateline or that we never see them. Some noise is expected and the sample sizes mean this can be averaged out.
David Bennett	East Midlands Trains	2 April 2018	At Nottingham, only one of the two entrances/exits is gated, so potentially not suitable.	We have considered replacing Nottingham with **Leeds**, **Sheffield, Derby or Leicester**, depending on gateline data availability from TOCs and subject to discussion with the Department for Transport. We have finally decided to include Derby following feedback from the stakeholders.
Robert Fickling	Rail North/ Transport for the North	3 April 2018	Concerns about LENNON validation given the gap on Oyster/PTE/multi-modal products, suggest using passenger counts instead.	We agree with the fact that there are gaps in LENNON for those products. **Oyster/PTE/multi-modal** products, however, have Journeys per Ticket estimates **based on surveys** or other data and do not rely upon the LENNON's Journeys per Ticket estimates on which this study is focused. Travel (revenue and associated journeys) on PTE type ticket are often also entered into LENNON as HQ items. When we select the 20 flows from each station, we will seek to remove those flows with a high proportion of passengers using those PTE type products, this is necessary to allow to compare information from gateline data with sales data LENNON. Passenger **counts do not indicate** the **ticket type or flow** so would not be suitable for validating gateline data as per our methodology as we will be investigating highly specified flows.
			Station selection does not reflect the issue of the daily to Season ticket price ratio.	The daily to Season **ticket price ratio** issue will be reflected in the **choice** of the **flows** for each station, where we will selection a range of destinations covering various price ticket ratios.

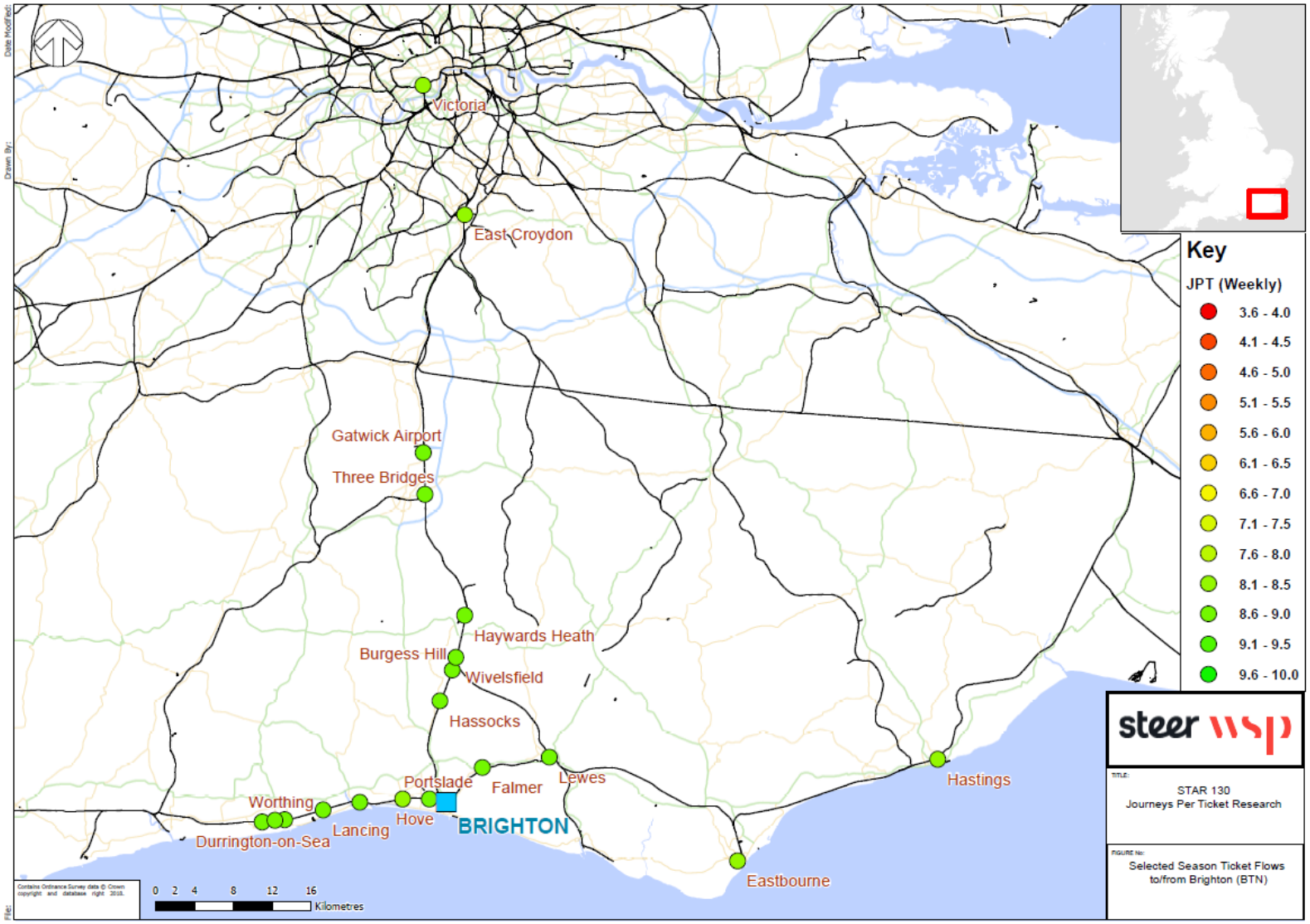
Author	Organisation	Date Received	Feedback	Our Response
			There are no big attractor stations represented in the north (other than Newcastle, with limited commuting). Leeds, Liverpool Central, Manchester Victoria or Manchester Piccadilly could be added to the station selection.	The listed stations have all be considered, however, we have identified **issues** with respect to these **stations**: - Leeds gateline is not managed by Cubic. Members of the project team have experience of requesting gateline data from Leeds in the past for other work without success but we will request it for this study; - Liverpool Central is not gated; - All Manchester termini have the same issue as the London termini: there are tickets to all Manchester termini, so difficult to compare gateline data to LENNON sales. **We will, however, capture demand to/from these stations through analysis of stations at the opposite end of the flows for example from Huddersfield or Chester.**
				The distance ranges included were **indicative** of potential commuter flows from those stations. These will be refined once the flows to be analysed for each station are selected.
James Angus	Network Rail	3 April 2018	There is no segmentation between in-boundary and out-boundary around London.	This research is focused on updating the standard Journeys per Ticket estimates in LENNON; Oyster/PTE/multi-modal products calculate Journeys per Ticket estimates are usually based on surveys or similar and do not rely upon LENNON's Journeys per Ticket estimates. By design we are not including in-boundary flows where Oyster is the dominant product used. Evidence such as Oyster Clicks is better suited to understanding in-boundary travel patterns and Oyster cards include a UID which can be used by TfL to directly understand Journeys per Ticket estimates for Oyster products.
			London and London Travelcard Area flows are underrepresented.	Our revised list of stations includes two London termini: Marylebone and Liverpool Street, of a total of 15 stations. As we are excluding in-boundary/LTA flows, we consider that these two London termini provide a **good representation of out-boundary commuting into/out of London**. We agree that selecting Charing Cross would not allow validating with LENNON due to the issue described. Therefore, we suggest using **Liverpool Street to replace Charing Cross**.
			Issue with passengers with tickets valid to multiple London termini, e.g. Charing Cross.	Although tickets to Liverpool Street would also be London Terminals tickets, there would only be a few interchange points, e.g. Stratford or Tottenham Hale. It would not be financially sensible for passengers to buy a London Terminals ticket and then interchange in Stratford as opposed to buying a Stratford ticket and then using the tube, therefore we consider using Liverpool Street we should be able to make a clean comparison of gateline data vs LENNON for London Terminals tickets.

Appendix B: Station and Flow Maps

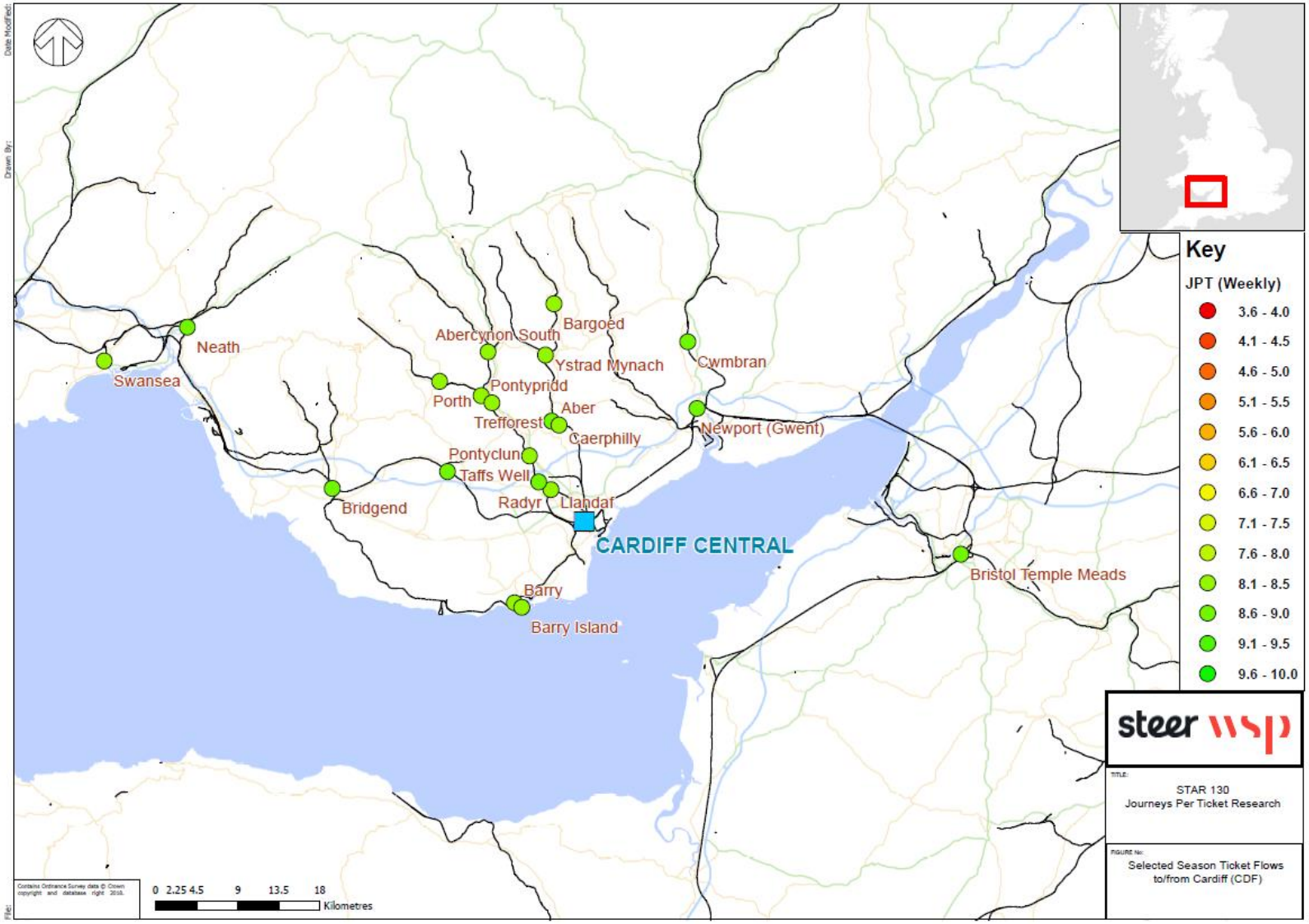
This appendix contains maps of the 15 stations and the flows we have used as part of our analysis of Journeys per Ticket estimates.

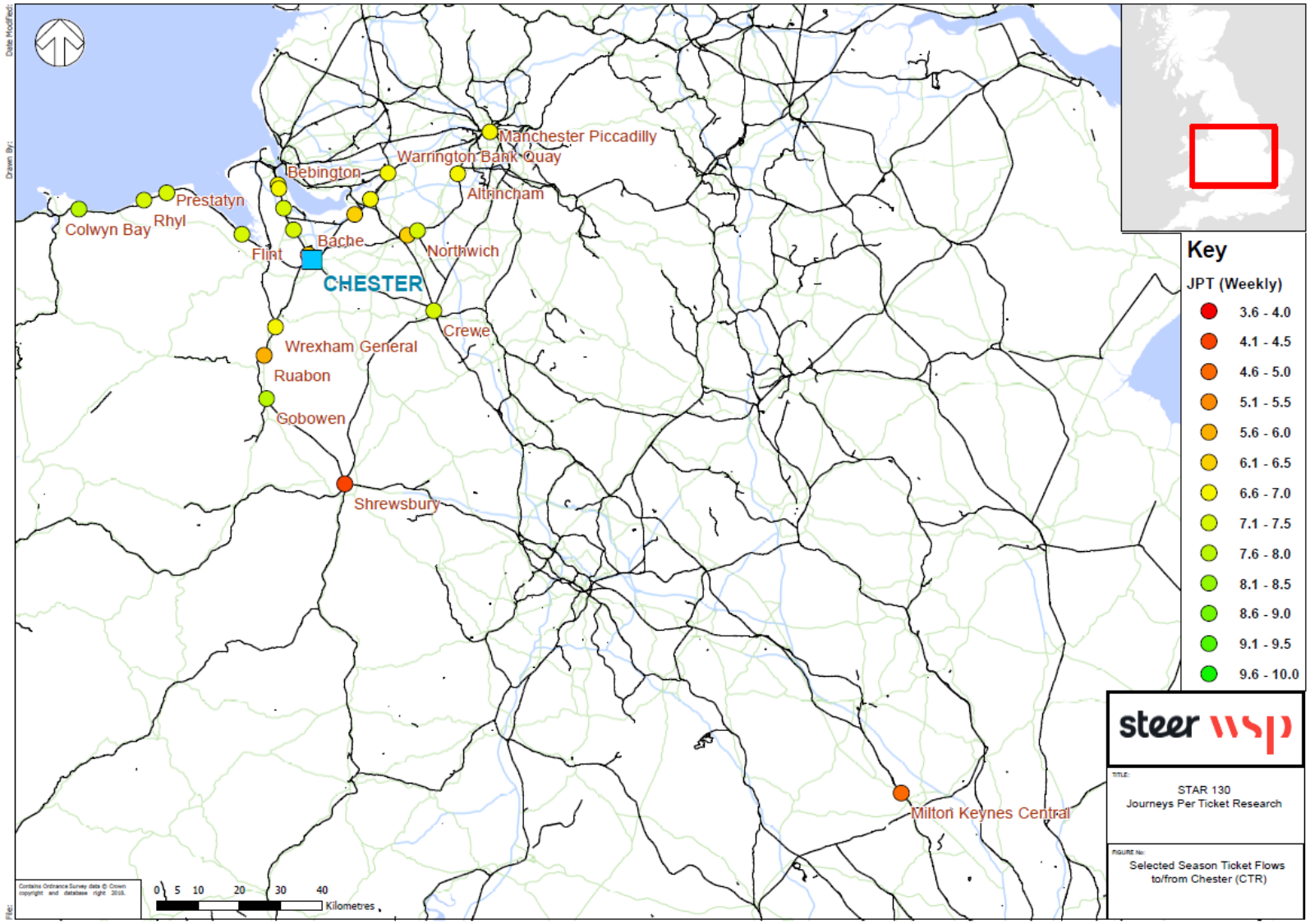
As an aid to understand the background of the following maps, the dark lines in the background represent the railway lines comprising the GB rail network, and the clear lines in the background represent the surrounding road network.

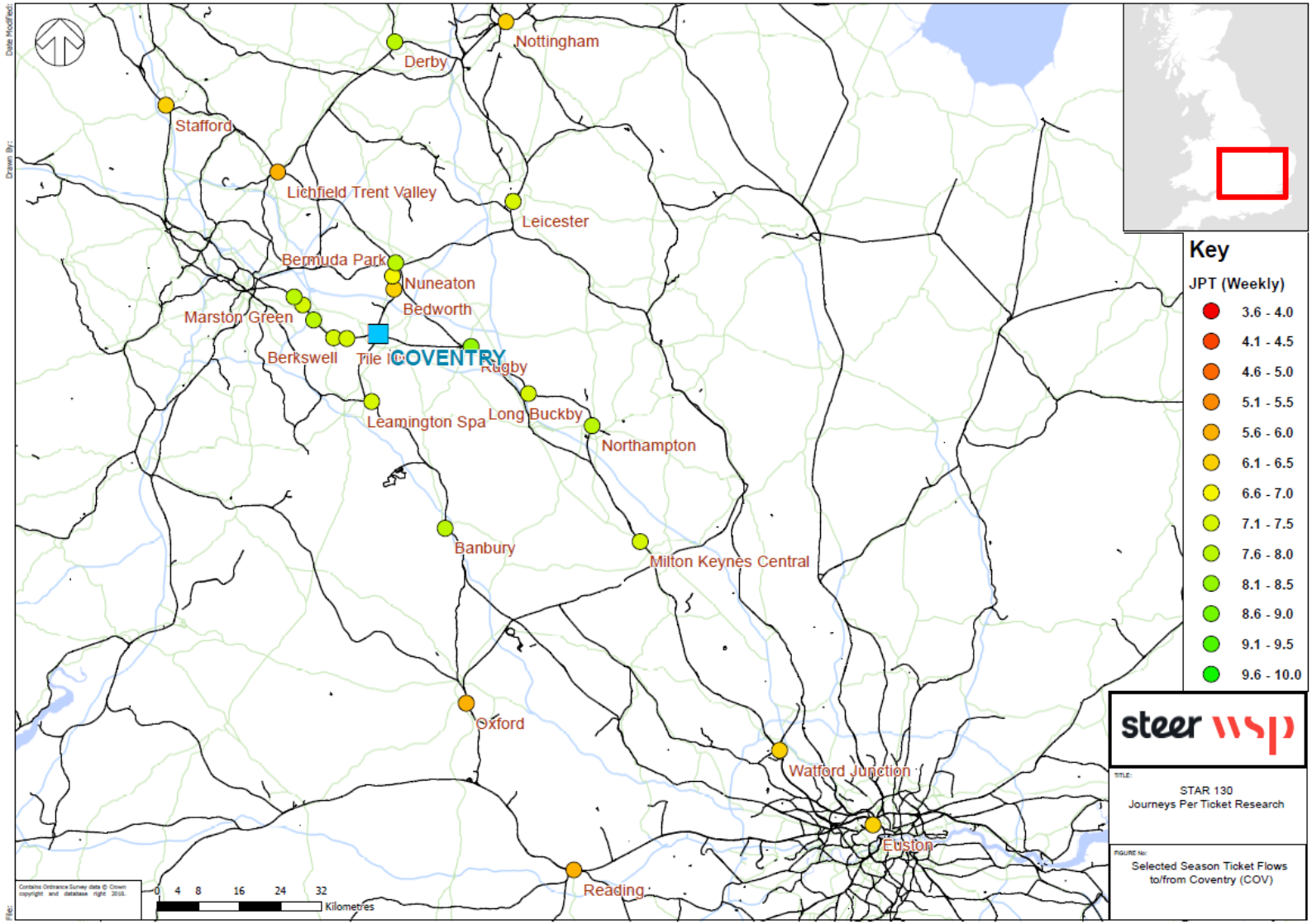

Oxford
Slough
Waterloo
Clapham Junction
Wimbledon
East Croydon
Reading
Bramley (Hampshire)
BASINGSTOKE
Hook
Fleet
Farnborough (Main)
Woking
Guildford
Whitchurch (Hampshire)
Overton
Andover
Salisbury
Winchester
Eastleigh
Southampton Central
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows
to/from Basingstoke (BSK)
0 2.254.5 9 13.5 18
Kilometres


Victoria
East Croydon
Gatwick Airport
Three Bridges
Haywards Heath
Burgess Hill
Wivelsfield
Hassocks
Portslade
Falmer
Lewes
Hastings
Worthing
Hove
Lancing
BRIGHTON
Durrington-on-Sea
Eastbourne
0 2 4 8 12 16
Kilometres
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows
to/from Brighton (BTN)

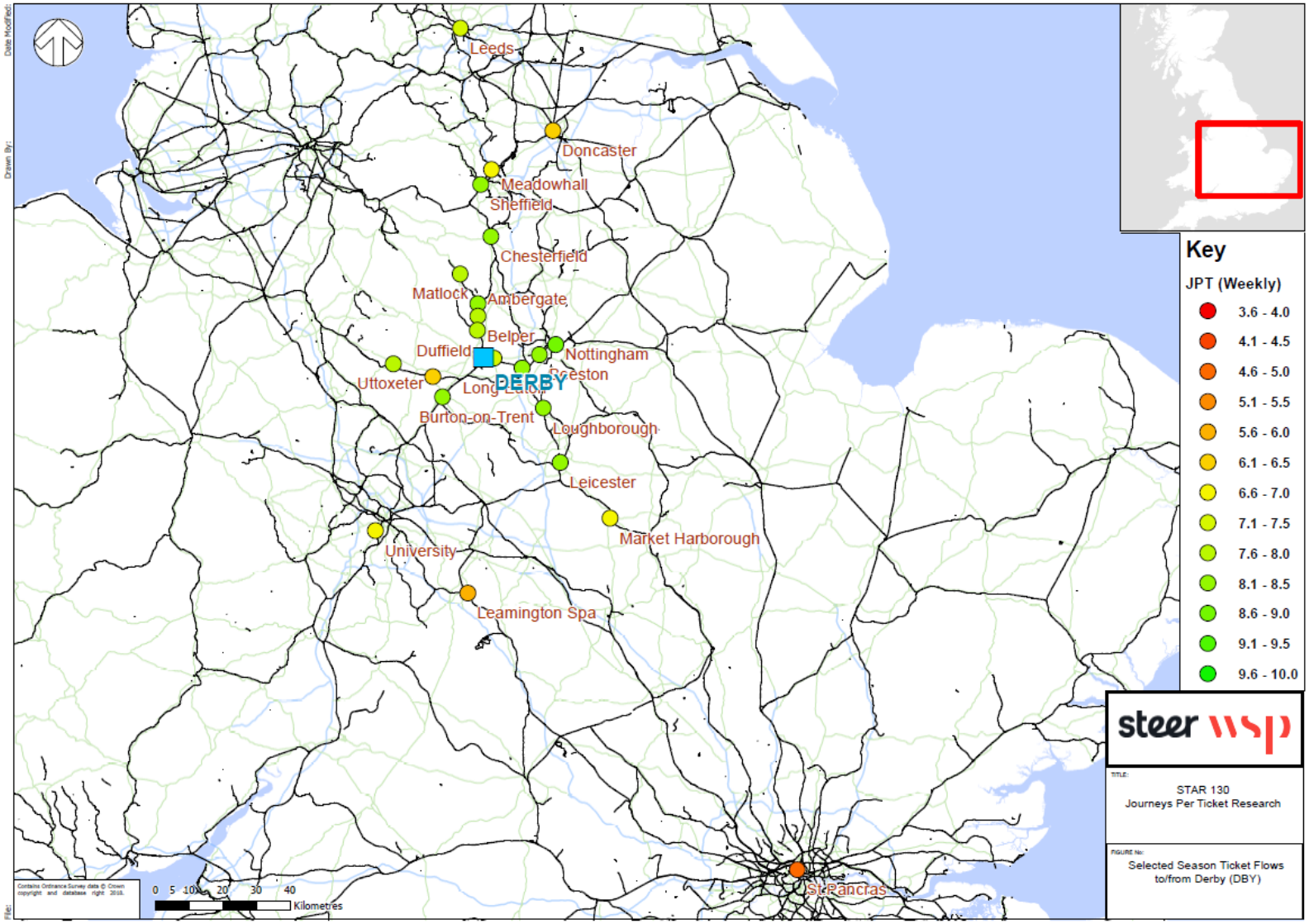
King's Lynn
Downham Market
Peterborough
March
Littleport
Ely
Waterbeach
Newmarket
CAMBRIDGE
Shepreth
Shelford
Whittlesford
Meldreth
Great Chesterford
Royston
Letchworth
Audley End
Baldock
Newport (Essex)
Hitchin
Stevenage
Bishops Stortford
Liverpool Street
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
0 3.5 7 14 21 28
Kilometres
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows
to/from Cambridge (CBG)

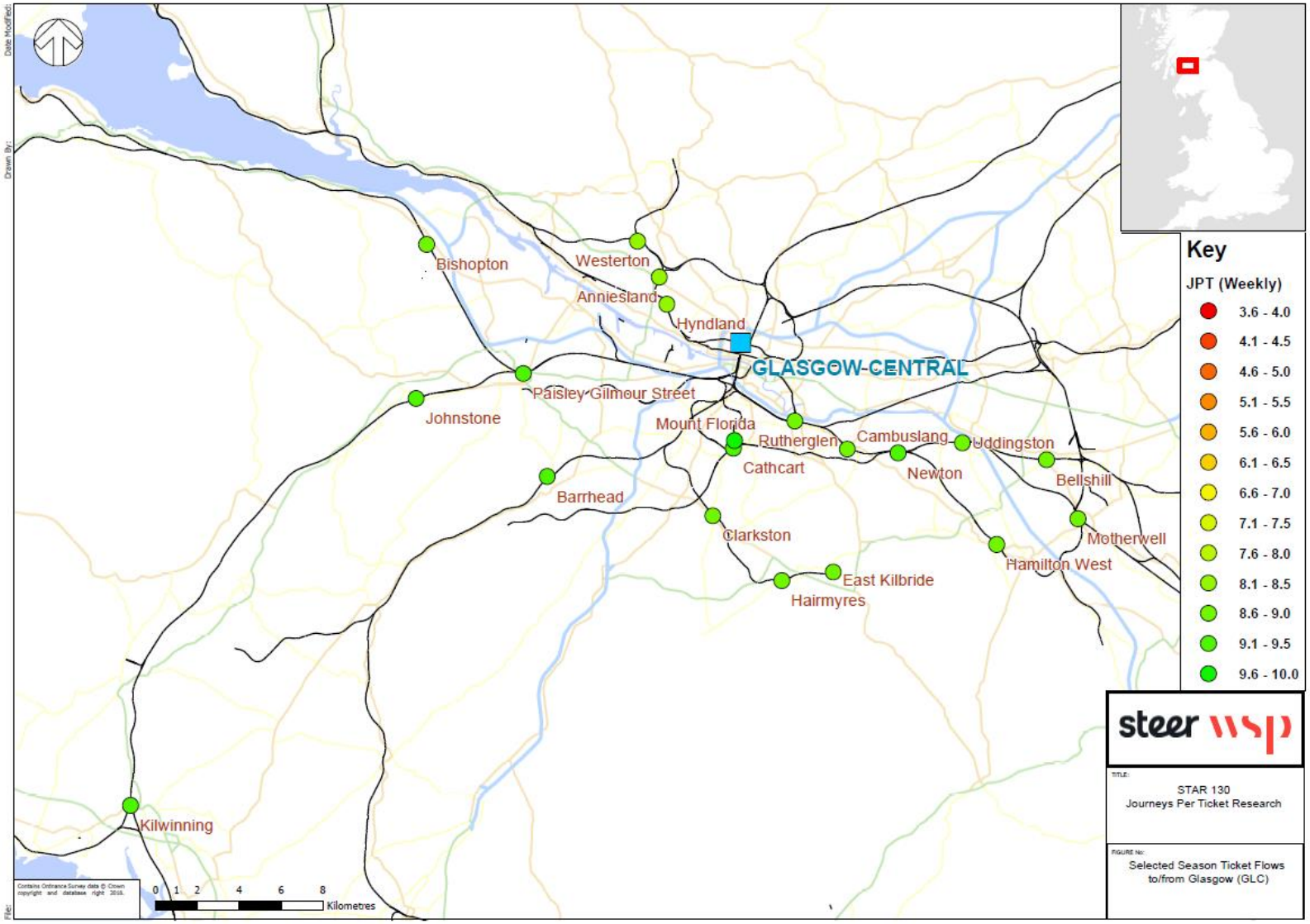

Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
Swansea
Neath
Bridgend
Abercynon South
Bargoed
Ystrad Mynach
Pontypridd
Porth
Aber
Treforest
Caerphilly
Pontyclun
Taffs Well
Radyr
Llandaf
CARDIFF CENTRAL
Barry
Barry Island
Cwmbran
Newport (Gwent)
Bristol Temple Meads
0 2.25 4.5 9 13.5 18
Kilometres
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows
to/from Cardiff (CDF)

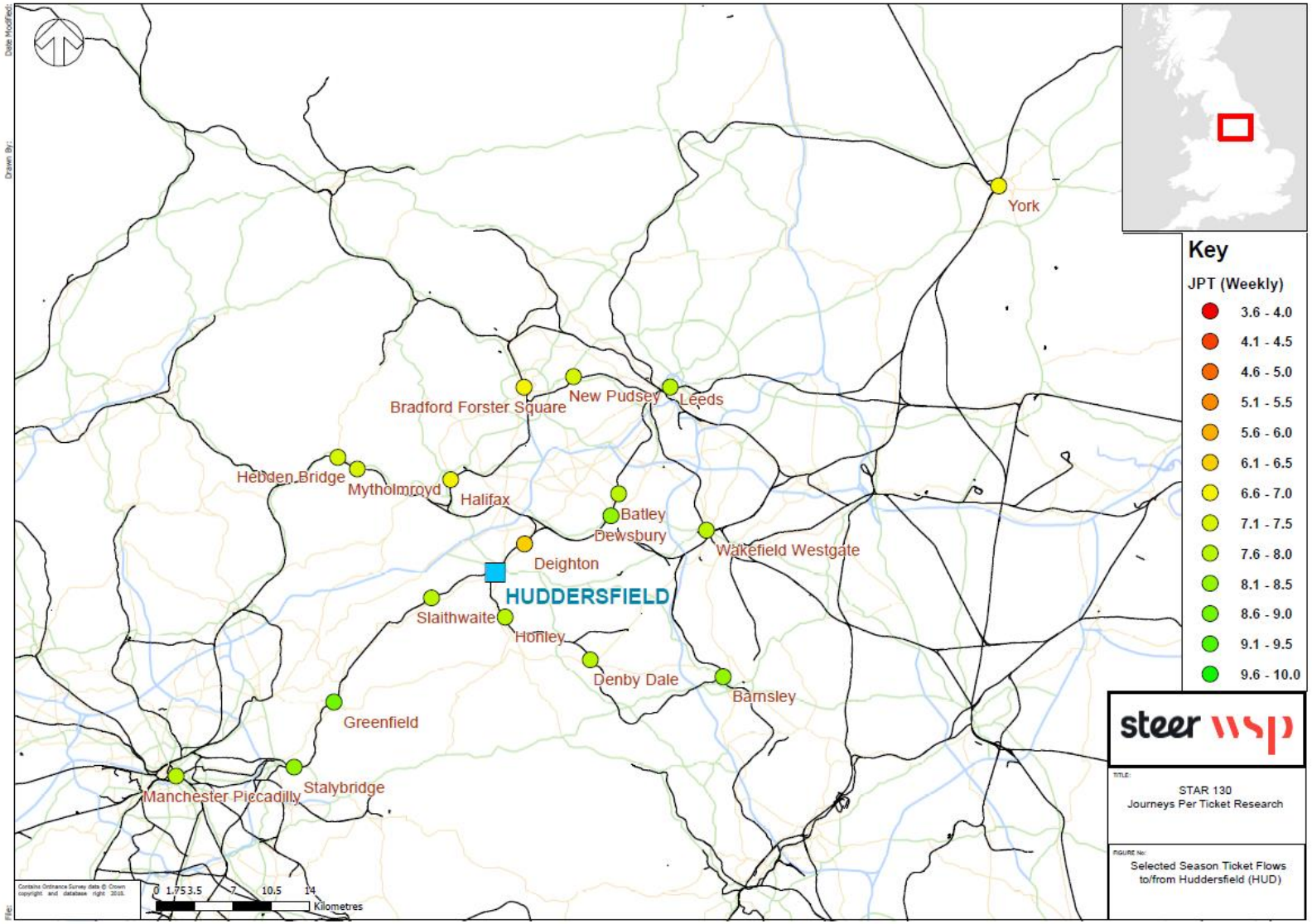

Manchester Piccadilly
Warrington Bank Quay
Bebington
Altrincham
Prestatyn
Rhyl
Colwyn Bay
Flint
Bache
Northwich
CHESTER
Crewe
Wrexham General
Ruabon
Gobowen
Shrewsbury
Milton Keynes Central
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Chester (CTR)
0 5 10 20 30 40
Kilometres


Nottingham
Derby
Stafford
Lichfield Trent Valley
Leicester
Bermuda Park
Nuneaton
Bedworth
Marston Green
Berkswell
Tile Hill
COVENTRY
Rugby
Leamington Spa
Long Buckby
Northampton
Banbury
Milton Keynes Central
Oxford
Watford Junction
Euston
Reading
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Coventry (COV)
0 4 8 16 24 32
Kilometres

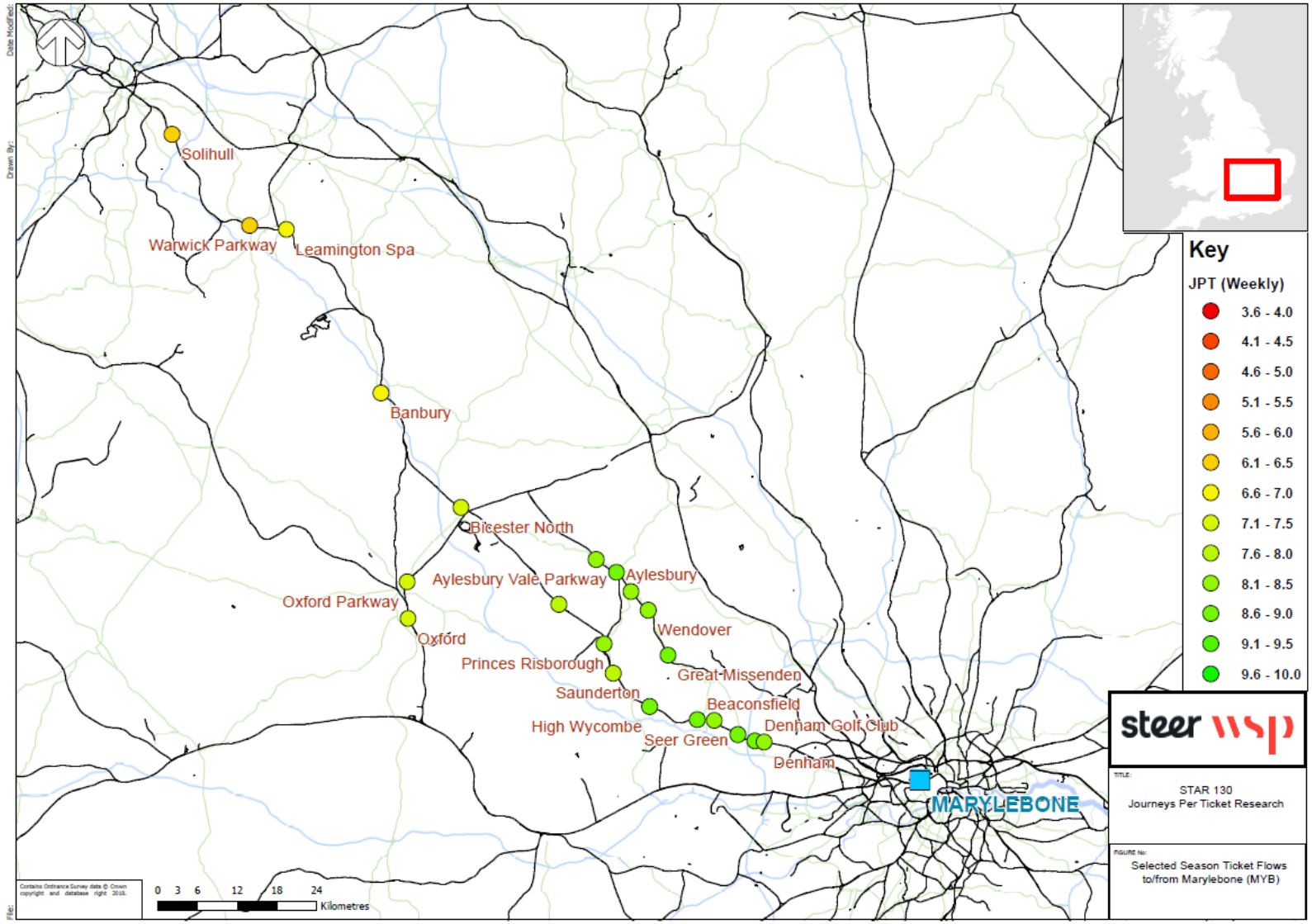
Manchester Piccadilly
Warrington Bank Quay
Liverpool Lime Street
Liverpool South Parkway
Runcorn
Stockport
Wilmslow
Hartford
Chester
Holmes Chapel
Sandbach
CREWE
Nantwich
Stoke-on-Trent
Stafford
Wolverhampton
Birmingham New Street
Milton Keynes Central
Euston
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
0 5 10 20 30 40
Kilometres
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows
to/from Crewe (CRE)

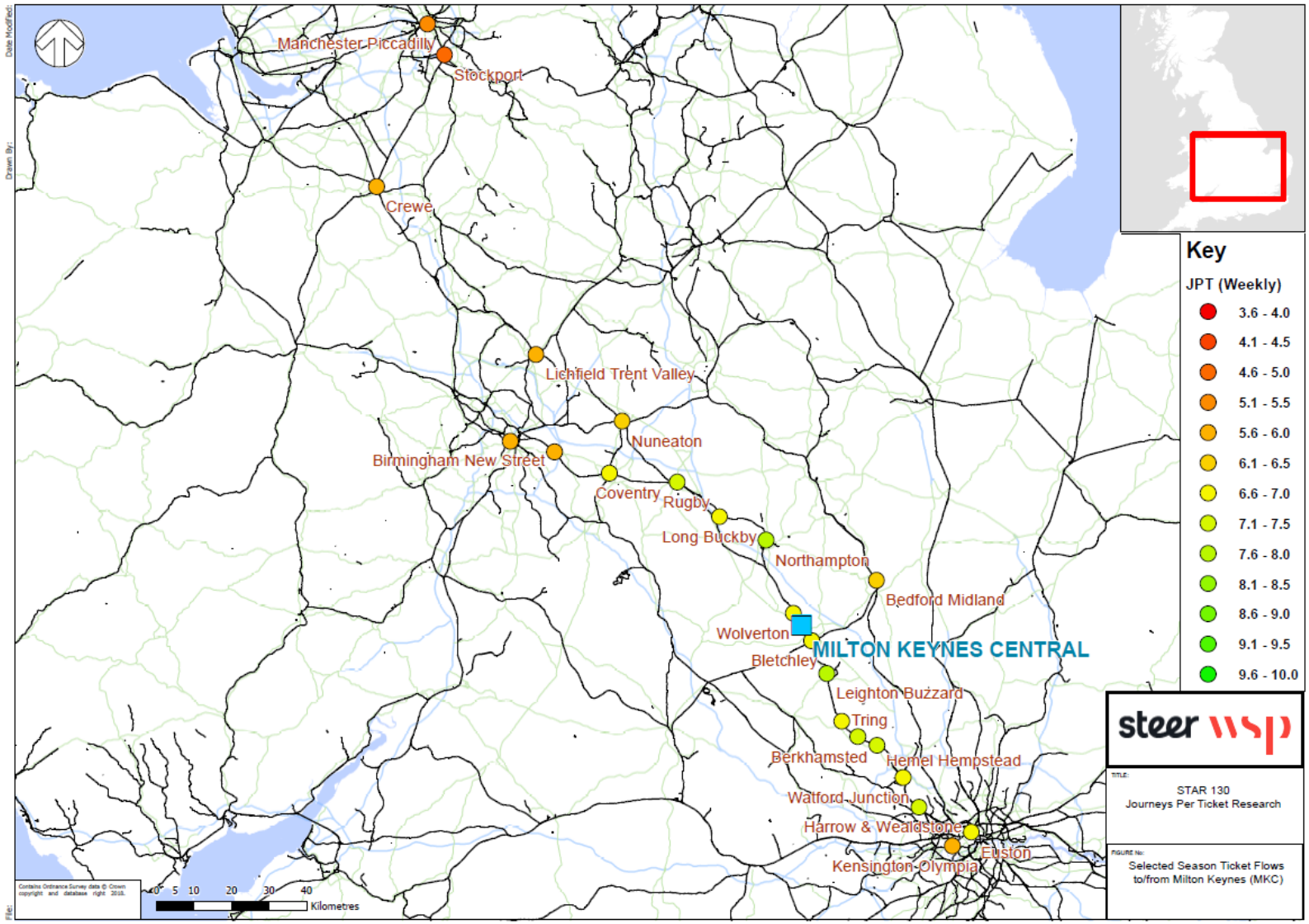

Leeds
Doncaster
Meadowhall
Sheffield
Chesterfield
Matlock
Ambergate
Belper
Duffield
Nottingham
Beeston
Uttoxeter
DERBY
Burton-on-Trent
Loughborough
Leicester
Market Harborough
University
Leamington Spa
St Pancras
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
0 5 10 20 30 40
Kilometres
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows
to/from Derby (DBY)

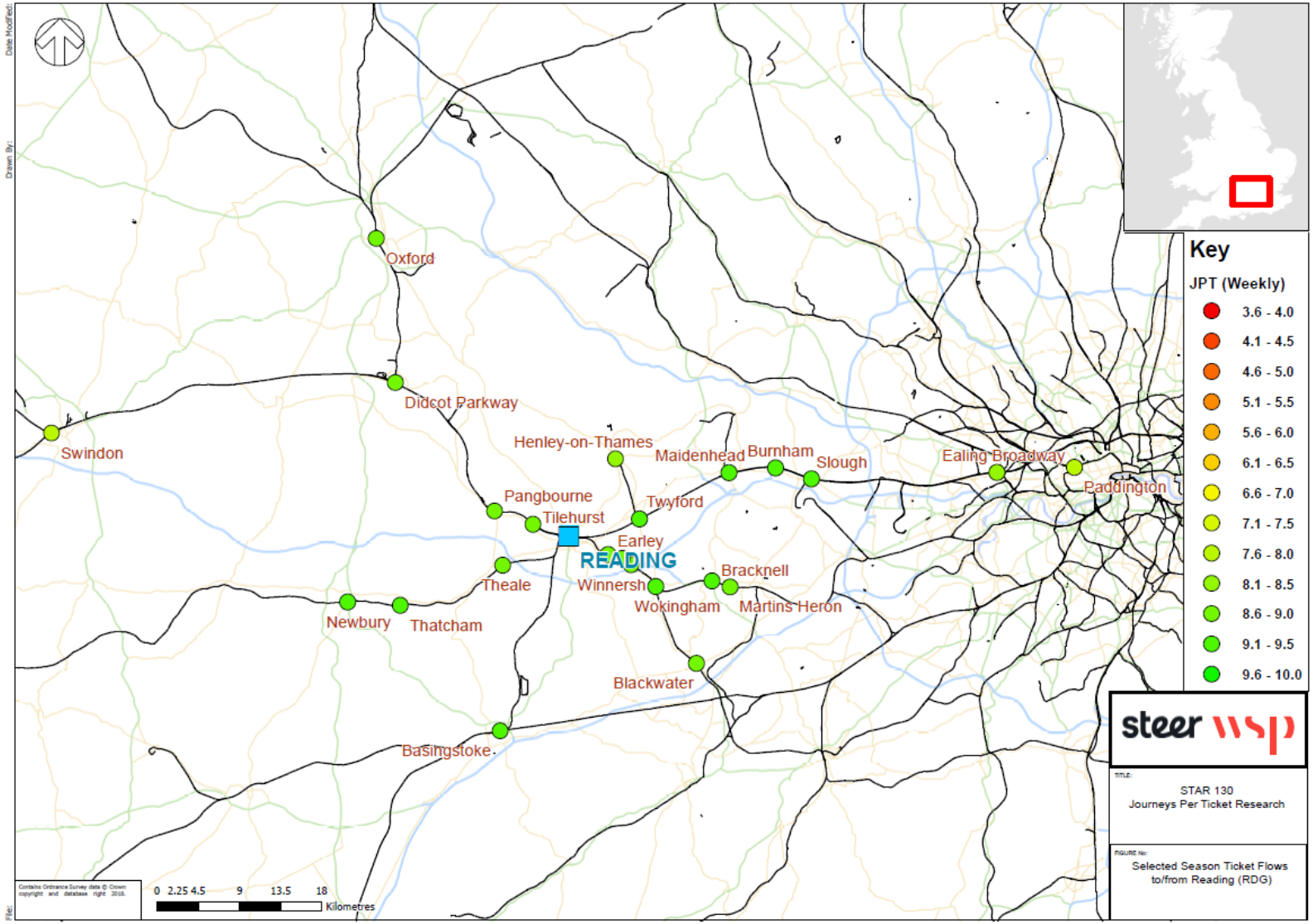

Bishopton
Westerton
Anniesland
Hyndland
GLASGOW-CENTRAL
Paisley Gilmour Street
Johnstone
Mount Florida
Rutherglen
Cambuslang
Uddingston
Cathcart
Newton
Bellshill
Barrhead
Clarkston
Motherwell
Hamilton West
East Kilbride
Hairmyres
Kilwinning
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Glasgow (GLC)
0 1 2 4 6 8
Kilometres


York
Bradford Forster Square
New Pudsey
Leeds
Hebden Bridge
Mytholmroyd
Halifax
Batley
Dewsbury
Wakefield Westgate
Deighton
HUDDERSFIELD
Slaithwaite
Honley
Denby Dale
Barnsley
Greenfield
Stalybridge
Manchester Piccadilly
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Huddersfield (HUD)
0 1.75 3.5 7 10.5 14
Kilometres

Ipswich
Manningtree
Colchester
Kelvedon
Witham
Chelmsford
Bishops Stortford
Harlow Town
Broxbourne
Cheshunt
Chingford
Highams Park
Harold Wood
Brentwood
Ingatestone
Shenfield
Woodham Ferrers
Wickford
Southend Victoria
Gidea Park
Goodmayes
LIVERPOOL STREET
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Liverpool Street (LST)
0 2 4 8 12 16
Kilometres


Solihull
Warwick Parkway
Leamington Spa
Banbury
Bicester North
Aylesbury Vale Parkway
Aylesbury
Oxford Parkway
Oxford
Wendover
Princes Risborough
Great Missenden
Saunderton
Beaconsfield
High Wycombe
Denham Golf Club
Seer Green
Denham
MARYLEBONE
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Marylebone (MYB)
0 3 6 12 18 24
Kilometres


Manchester Piccadilly
Stockport
Crewe
Lichfield Trent Valley
Nuneaton
Birmingham New Street
Coventry
Rugby
Long Buckby
Northampton
Bedford Midland
Wolverton
MILTON KEYNES CENTRAL
Bletchley
Leighton Buzzard
Tring
Berkhamsted
Hemel Hempstead
Watford Junction
Harrow & Wealdstone
Euston
Kensington Olympia
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Milton Keynes (MKC)
0 5 10 20 30 40
Kilometres


Oxford
Didcot Parkway
Swindon
Henley-on-Thames
Maidenhead
Burnham
Slough
Ealing Broadway
Paddington
Pangbourne
Tilehurst
Twyford
Earley
READING
Theale
Winnersh
Bracknell
Wokingham
Martins Heron
Newbury
Thatcham
Blackwater
Basingstoke
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
steer wsp
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Reading (RDG)
0 2.25 4.5 9 13.5 18
Kilometres

Carmarthen
Ammanford
Ferryside
Kidwelly
Llangennech
Llansamlet
Skewen
Pembrey & Burry Port
Llanelli
Neath
Gowerton
SWANSEA
Baglan
Port Talbot Parkway
Pyle
Pencoed
Pontyclun
Bridgend
Newport (Gwent)
Cathays
Cardiff Queen Street
Cardiff Bay
Rhoose (Cardiff Int Airpt)
Key
JPT (Weekly)
3.6 - 4.0
4.1 - 4.5
4.6 - 5.0
5.1 - 5.5
5.6 - 6.0
6.1 - 6.5
6.6 - 7.0
7.1 - 7.5
7.6 - 8.0
8.1 - 8.5
8.6 - 9.0
9.1 - 9.5
9.6 - 10.0
0 2.25 4.5 9 13.5 18
Kilometres
STAR 130
Journeys Per Ticket Research
Selected Season Ticket Flows to/from Swansea (SWA)

Control Information

Prepared by

Steer
28-32 Upper Ground
London SE1 9PD
+44 20 7910 5000
www.steergroup.com

Prepared for

Department for Transport
Great Minster House
33 Horseferry Road
Westminster
London
SW1P 4DR

Steer project/proposal number

23257401

Client contract/project number

PPRO 04 101 130

Author/originator

Antonio Beltrán Arranz

Reviewer/approver

Andrew Davies

Other contributors

Mark Brownie
Nick King
Julian Phatarfod

Distribution

Client: Department for Transport

Steer: Project Team

Version control/issue number	Date
v1.00 – Draft Report Issue	December 2018
v2.00 – Revised Draft Report with responses to client comments	January 2019
v3.00 – Final Report Issue	February 2019
v4.00 – Final Report Issue with final comments incorporated	February 2019
v5.00 – Final Report for publication	April 2019

steergroup.com