Department for Transport

Revisiting the Elasticity Based Framework - Update

Final Report

PPRO 04/73/03

Rev A | 15 February 2012

Ove Arup & Partners Ltd
13 Fitzroy Street
London
W1T 4BQ
United Kingdom
www arup com

Job number 207999-10

ARUP

Contents

Tables

Figures

Appendices

Executive Summary

This document reports on the work carried out by Arup and Oxera to update the data and outputs from the 'Revisiting the Elasticity Based Framework' (RTEBF) project. The work was divided into twelve tasks. The main report gives a brief overview of the work carried out for each task. Where necessary, this is supplemented by a technical appendix.

If further details are required on any of the tasks, please refer to the table below for whom to contact in the first instance, although it is recommended that both Arup and Oxera are copied into any correspondence.

Table E. 1: Tasks and Organisation

Task number	Task title	Contact Organisation
1	Reconciling car ownership data with raw NTEM data	Arup
2	Check market segmentation fields in TOAD prepared by DfT	Arup
3	Resolve issues identified in tasks 1 and 2	Arup
4	Include service quality index	Oxera
5	Update PPM field	Arup
6	Provide Full Meta-Data Within TOAD Access Database File	Arup
7	Review and Populate SDG 'TOAD Summary Data' Sheet	Arup
8	Provide Full Description and Schematic Diagram of Stata Files	Oxera
9	Update Guidance and Findings Reports as Necessary	Oxera
10	Repeat General to Specific Modelling on Three Market Segments	Oxera
11	Review Relevant Worksheets in RAFF Model	Arup
12	Ticket Aggregation and Car Journey Times	Oxera

1 Reconciling Car Ownership Data with Raw NTEM Data

1.1 Introduction

The DfT requires clarification that the car ownership data used in the TOAD analysis is consistent with the publically available NTEM data. This is to ensure that future updates of NTEM data published via TEMPRO can be integrated into the analysis if necessary.

This section outlines the reconciliation undertaken by Arup to confirm that the car ownership data used in TOAD is consistent with both the source data issued by the DfT and the NTEM data.

1.2 Reconciliation of the data

DfT provided car ownership (as well as population and employment) data for the following years in December 2008 via the Livelink project extranet:

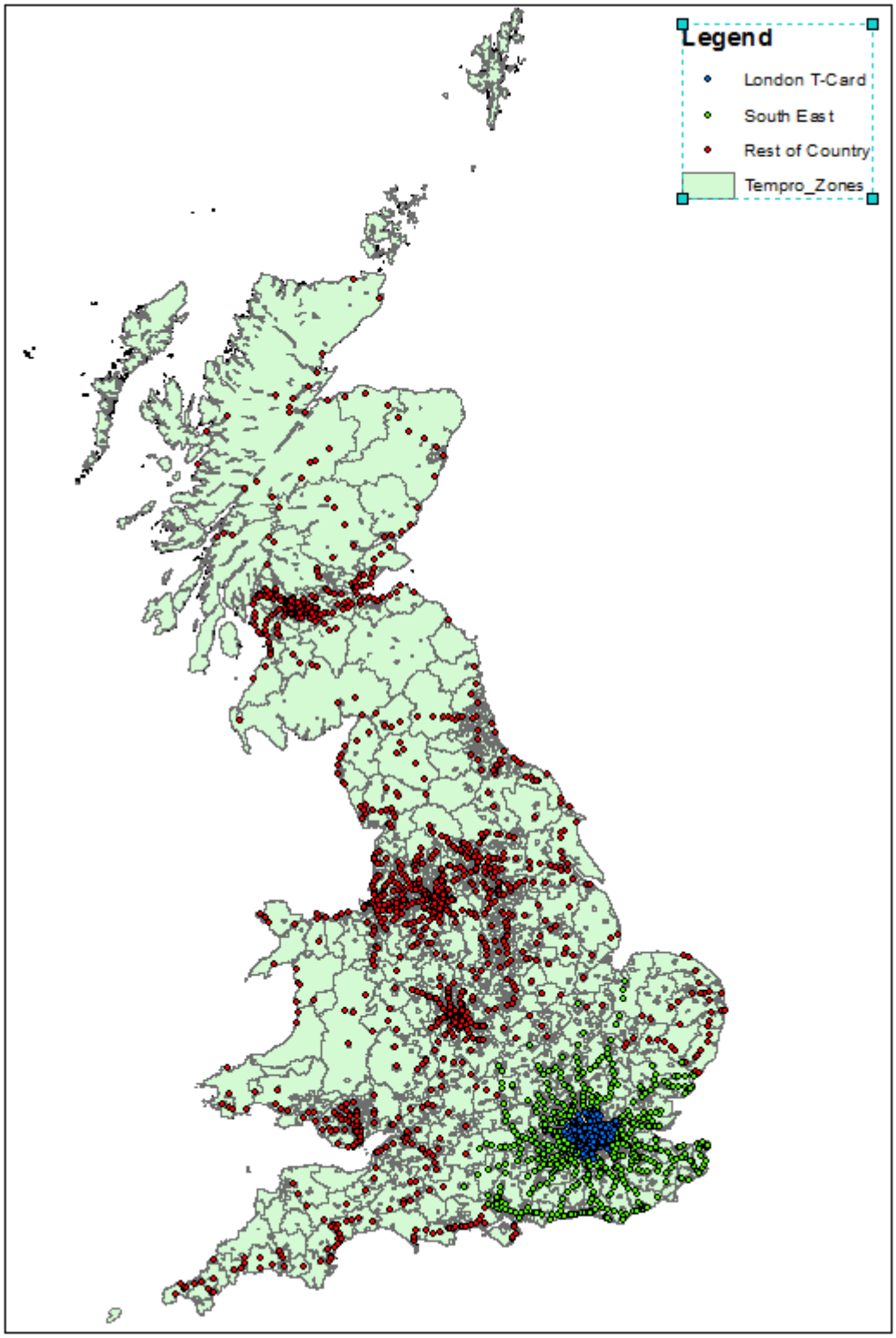
- 1991;
- 2001;
- 2006; and
- 2011.

This data was compiled into a single spreadsheet called TEMPRO ALL YEARS.xls. For years between those provided, a linear interpolation was used to estimate car ownership values. A sample of the car ownership data in this spreadsheet has been compared to NTEM 5.4 data extracted using TEMPRO v6.0 and has been found to be an exact match.

For TOAD, the origin-destination station pairs require car ownership data at the origin. This was achieved using GIS, by overlaying the UK rail station network on top of the NTEM zonal structure as shown in

Figure **1.1** below. The NTEM zone for each station was then recorded in a spreadsheet and the NTEM car ownership data for the zone was associated with each relevant station. For each O-D pair in TOAD, by year, the relevant origin car ownership data was extracted from this spreadsheet. A sample of the car ownership data in TOAD has been compared to the TEMPRO ALL YEARS.xls spreadsheet and has been found to be an exact match.

Figure 1.1: Stations within TEMPRO Zones



It is therefore determined that the car ownership data in TOAD does reconcile with both the source data and the published NTEM 5.4 data, and that the correct NTEM data is applied to the relevant stations in TOAD.

1.3 Updating to NTEM 6.2

It is noted that the current version of NTEM data is 6.2, and it may be worthwhile, as an additional piece of work outside the scope of this RTEBF update, to revise the car ownership data in TOAD to reconcile with this the latest version. However, this would also require, for internal consistency in the modelling, updating the NTEM population and employment data, as well as potentially revising the car cost and car journey time variables. This would be a significant undertaking.

2 Check Market Segmentation Fields in TOAD Prepared by DfT

2.1 Introduction

Following the completion of the main study, DfT added market segmentation fields to the main TOAD database in order to facilitate further analysis. This work was carried out by a summer student, and Arup was asked to check that the fields were correctly added.

2.2 TOAD Database and Queries

We were supplied with the following Access database, which we understand was used to add the market segmentation variables:

- 100729_-_Toad_Summary_Data_(procedure).mdb

No additional documentation was available, however, there were eight update queries in the database, which appear to create the variables Originsegment and Destinationsegment. These are reproduced below.

We have reviewed these queries, as well as a random sample of 50 flows in the database, and we are satisfied that they correctly represent the market segmentation used in the main analysis.

Table 2.1: TOAD Database Queries

Query Name	SQL
Origin Airport	UPDATE Toad SET OriginSegment = "Airport" WHERE originname="GATWICK AIRPORT" OR originname="MANCHESTER AIRPORT" OR originname="BIRMINGHAM INTERNATIONAL" OR originname="SOUTHAMPTON AIRPORT PARKWAY" OR originname="STANSTED AIRPORT" OR originname="LUTON AIRPORT PARKWAY";
Origin Core	UPDATE Toad SET OriginSegment = "Core" WHERE (originname='BIRMINGHAM BR' Or originname='MANCHESTER BR' Or originname='LIVERPOOL' Or originname='NOTTINGHAM' Or originname='BRISTOL TEMPLE MEADS' Or originname='BRISTOL PARKWAY' Or originname='SHEFFIELD' Or originname='CARDIFF BR' Or originname='EDINBURGH' Or originname='GLASGOW BR' Or originname='NEWCASTLE' Or originname='LEEDS' Or originname='LEICESTER' Or originname='YORK' Or originname='HULL');
Origin LSEE	UPDATE Toad SET OriginSegment = "LSEE" WHERE (O_gor="London" Or O_gor="South East" Or O_gor="East of England");
Origin Other	UPDATE Toad SET OriginSegment = "Other" WHERE (originname<>'BIRMINGHAM BR') And (originname<>'MANCHESTER BR') And (originname<>'LIVERPOOL') And (originname<>'NOTTINGHAM') And (originname<>'BRISTOL TEMPLE MEADS') And (originname<>'BRISTOL PARKWAY') And (originname<>'SHEFFIELD') And (originname<>'CARDIFF BR') And (originname<>'EDINBURGH') And (originname<>'GLASGOW BR') And (originname<>'NEWCASTLE') And (originname<>'LEEDS') And (originname<>'LEICESTER') And (originname<>'YORK') And (originname<>'HULL') And (O_gor<>'South East') And (O_gor<>'London') And (O_gor<>'East of England') And (originname<>'GATWICK AIRPORT') And (originname<>'MANCHESTER AIRPORT') And (originname<>'BIRMINGHAM INTERNATIONAL') And (originname<>'SOUTHAMPTON AIRPORT PARKWAY') And (originname<>'STANSTED AIRPORT') And (originname<>'LUTON AIRPORT PARKWAY');
Destination Airport	UPDATE Toad SET destinationSegment = "Airport" WHERE destinationname="GATWICK AIRPORT" OR destinationname="MANCHESTER AIRPORT" OR destinationname="BIRMINGHAM INTERNATIONAL" OR destinationname="SOUTHAMPTON AIRPORT PARKWAY" OR destinationname="STANSTED AIRPORT" OR destinationname="LUTON AIRPORT PARKWAY";
Destination Core	UPDATE Toad SET DestinationSegment = "Core" WHERE (destinationname='BIRMINGHAM BR' Or destinationname='MANCHESTER BR' Or destinationname='LIVERPOOL' Or destinationname='NOTTINGHAM' Or destinationname='BRISTOL TEMPLE MEADS' Or destinationname='BRISTOL PARKWAY' Or destinationname='SHEFFIELD' Or destinationname='CARDIFF BR' Or destinationname='EDINBURGH' Or destinationname='GLASGOW BR' Or destinationname='NEWCASTLE' Or destinationname='LEEDS' Or destinationname='LEICESTER' Or destinationname='YORK' Or destination

	name='HULL');
Destination LSEE	UPDATE Toad SET DestinationSegment = "LSEE" WHERE (D_gor="London" Or D_gor="South East" Or D_gor="East of England");
Destination Other	UPDATE Toad SET DestinationSegment = "Other" WHERE (destinationname<>'BIRMINGHAM BR') And (destinationname<>'MANCHESTER BR') And (destinationname<>'LIVERPOOL') And (destinationname<>'NOTTINGHAM') And (destinationname<>'BRISTOL TEMPLE MEADS') And (destinationname<>'BRISTOL PARKWAY') And (destinationname<>'SHEFFIELD') And (destinationname<>'CARDIFF BR') And (destinationname<>'EDINBURGH') And (destinationname<>'GLASGOW BR') And (destinationname<>'NEWCASTLE') And (destinationname<>'LEEDS') And (destinationname<>'LEICESTER') And (destinationname<>'YORK') And (destinationname<>'HULL') And (D_gor<>'South East') And (D_gor<>'London') And (D_gor<>'East of England') And (destinationname<>'GATWICK AIRPORT') And (destinationname<>'MANCHESTER AIRPORT') And (destinationname<>'BIRMINGHAM INTERNATIONAL') And (destinationname<>'SOUTHAMPTON AIRPORT PARKWAY') And (destinationname<>'STANSTED AIRPORT') And (destinationname<>'LUTON AIRPORT PARKWAY');

3 Resolve Issues Identified in Task 1

No issues were identified in tasks 1 and 2; therefore there is no requirement for this task.

4 Include Service Quality Index

There were two tasks in this update phase of the contract concerning updating the data. These involved the following:

- including a service quality index in the main dataset which varies by segment (task 4); and
- correcting an error in the PPM data matching (task 5), where the PPM data used in the original Revisiting study was offset by one year.

5 Update PPM Field

The flow performance dataset for TOAD has been updated to correct inconsistencies found in the original submission in 2010. The net result of updating these calculations is that there is a complete PPM dataset for all TOAD flows, for all years, with the correct data allocated to the correct years.

6 Provide Full Meta-Data Within TOAD Access Database File

This section provides advice on how to share the data more widely and in an effective way. This needs to happen to ensure that:

- version control issues are minimised;
- TOAD is accessible to all potential users;
- the data is maintained in a suitable format; and
- updates can be achieved at a minimum cost.

The following numbered items summarise ways of achieving the above aims:

1. **Master copy.** Retain the Stata version of TOAD as the master: all spreadsheet or Access versions should be considered secondary. This implies that all updates to TOAD should be done through statistical code in Stata, which will facilitate the continuity of an audit trail. This does not necessarily prevent copies of TOAD being distributed in Access, but they should be carefully controlled and, in the event of inconsistency arising, the Stata version should be considered the master;
2. **Owner.** It is suggested that an 'owner' of TOAD is nominated within DfT, who can control both internal and external distribution. There should be a succession plan for when that 'owner' moves to a new post;
3. **List.** It is suggested that a list is kept of people with copies of TOAD, who can be notified when the version of TOAD changes;
4. **Versions.** Ensure that an up to date list of changes between versions is maintained and kept with the master version;
5. **Readme file.** A readme file should be provided with the dataset, explaining what it is (and is not); and
6. **Original read-only.** Ensure that an original copy of TOAD is kept as a read-only file on the network, a copy of which is also saved off site.

It is also worth highlighting that, in order to maximise the value of this dataset, it should be advertised and made available for use both within and, subject to confidentiality restrictions, outside DfT. We also recommend that a regular update process to bring the dataset up to date may be useful to maximise the value of the asset, particularly given the substantial investment made in this study, of which TOAD is one of the major outputs.

7 Review and Populate SDG 'TOAD Summary Data' Sheet

We have reviewed and populated the SDG 'TOAD Summary Data' sheet. The purpose of this task was to ensure that the content of TOAD is clear and documented in a way that can be easily shared with other parties. The workbook

contains summary information about TOAD definitions, PPM, SQI and 3-year elasticities.

An excel copy of the data sheet is provided with this report:

- RTEBF TOAD Summary Sheet.xls

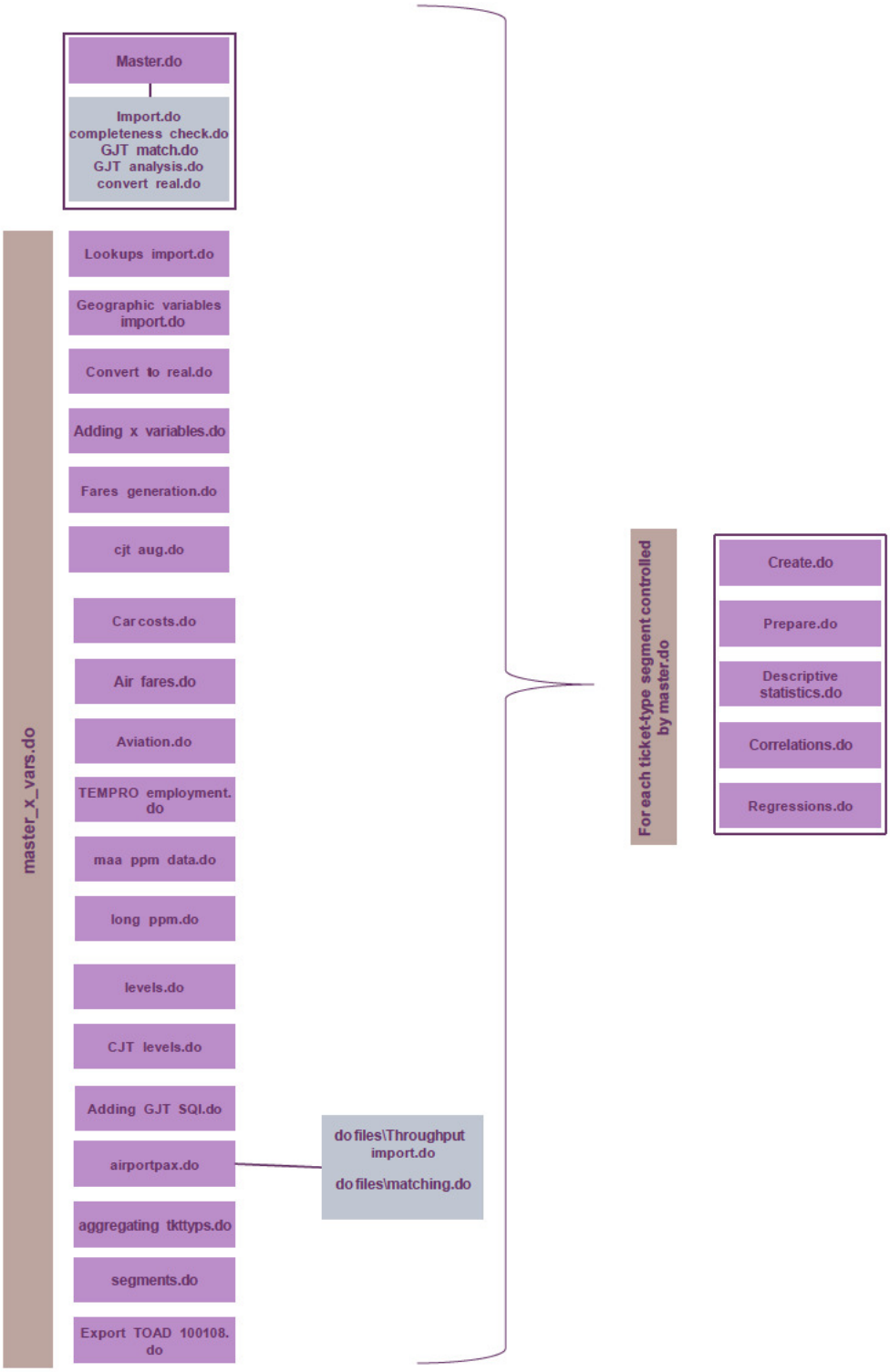
8 Provide Full Description and Schematic Diagram of Stata Files

Given the substantial volume of code created for this project, the diagrams are split into two as follows:

- a summary diagram of the service quality index creation, which illustrates the flow of the .do files (see **Figure 8.1**);
- a summary for the overall analysis, which illustrates the flow of the .do files (see **Figure 8.2**);

Please note that these structures reflect the revised .do files provide to the DfT. The original .do files are slightly different in some places, but are fundamentally unchanged.

Figure 8.1: Summary of overall .do files



Source: Oxera.

Figure 8.2: Summary of .do files

Source: Oxera.

9 Update *Guidance* and *Findings* Reports as Necessary

The Guidance and Findings reports have not been updated to reflect the changes to the LSEE-to-LSEE combined segment or the results of the analysis using the revised PPM variable. This is to ensure that the results contained within the Guidance and Findings reports are internally consistent and comparable.

Replacing only four of the segments with results which are based on updated PPM data would mean that the results in the reports for those segments would not be comparable with those in other segments.

In addition, although the LSEE-to-LSEE combined segment has changed, it would not be possible to identify which changes have arisen as part of the PPM change, and which have arisen as a result of the change in aggregation. Therefore, in order to avoid spurious comparisons being drawn, it is recommended that neither the Guidance report nor the Findings report is updated.

10 Repeat General to Specific Modelling on Three Market Segments

Revised econometric analysis has been undertaken to understand what the implications of the adjustments made to the PPM data were for the results of the econometric modelling.

The approach adopted by Oxera was to assume that the same specific econometric model is appropriate after the revision to the PPM data. This allows a like-for-like comparison between models which were produced in the original study and the models which were estimated using the revised PPM data. An alternative approach would have been to conduct a revised general to specific modelling procedure. However, this would have introduced the additional complication that, even if the dataset was identical, different modellers may produce different specific models. Of course, there is a possibility that the change in PPM data may have resulted in a different specific model being selected, but given the strength of the PPM variable as an explanatory factor, this seems unlikely.

Table 10.1, **Table 10.2** and **Table 10.3** summarise the differences between the models using the updated dataset, and those originally reported in the Revisiting reports. These segments were selected to provide a cross-section of the segments which were estimated for the Revisiting study. Full dashboards, for both the original and revised models are provided in Appendix A. These dashboards provide more detail, including the results of diagnostic tests, confidence intervals and a measure of model fit.

Table 10.1: Other to other, season

	One-year elasticities		**Three-year elasticities**	
	Revised	**Original**	**Revised**	**Original**
Fare	–0.97	–0.99	–1.33	–1.32
Income	-0.50	–0.58	–1.35	–1.50
Car cost	–0.42	–0.82	–0.36	–0.56
GJT	–0.04	–0.11	–0.67	–0.97
Performance	0.78	–0.23	3.97	2.68

Note: GJT stands for Generalised Journey Time. Numbers in bold lie outside the 95% confidence interval of the original results.

Source: Oxera analysis.

Table 10.2: Other to LSEE, full

	One-year elasticities		**Three-year elasticities**	
	Revised	**Original**	**Revised**	**Original**
Fare	–1.81	–1.81	–1.50	–1.5
Cross–price	0.38	0.38	0.51	0.51
Income	0.67	0.69	0.91	0.93
Population	1.88	1.88	2.54	2.53
Car cost	0.69	0.68	0.94	0.91
GJT	–0.15	–0.16	–1.17	–1.18
Service quality index	1.16	1.15	1.57	1.55

Note: GJT stands for Generalised Journey Time. Numbers in bold lie outside the 95% confidence interval of the original results.

Source: Oxera analysis.

Table 10.3: Core to LSEE, reduced

	One-year elasticities		Three-year elasticities	
	Revised	**Original**	**Revised**	**Original**
Fare	–0.85	–0.56	–1.06	–0.68
Income	–1.2	–0.55	**2.4**	1.4
Car cost	0.5	0.43	0.62	0.52
GJT	–0.44	–0.94	–0.56	–1.15
Performance	0.54	0.31	0.68	0.38

Note: GJT stands for Generalised Journey Time. Numbers in bold lie outside the 95% confidence interval of the original results.

Source: Oxera analysis.

As can be seen from the tables above, the revised elasticities are typically within the 95% confidence intervals of the original estimates and therefore, while the elasticity estimates have changed, in some cases noticeably, the difference would not be significant at the 5% level. A qualitative assessment suggests that the general effect of the PPM change is to increase the PPM elasticities, and reduce the GJT elasticities (in absolute terms), although the effect is variable across segments. The effect on the elasticities to other variables is changeable between segments. The fare elasticities are broadly unaffected by this change.

11 Review Relevant Worksheets in RAFF Model

The specific version of the spreadsheet that was reviewed by Arup was "RAFF_v0.82.xls", as implemented by SDG (Steer Davies Gleave). The spreadsheet was reviewed with particular focus on:

- the implementation of the formulae specified in the Guidance report;
- whether the expected results were produced in stress tests comparing the responses to specific percentage changes in each variable.

The spreadsheet is laid out with a range of differently coloured tabs reflecting user inputs containing the formulae (green), lookups containing correspondences and assumptions (yellow), base data and demand drivers with forecast growth (purple), and the model outputs (white). The spreadsheet has an input tab "I_Control" where all the model run parameters can be changed and the model can be run.

The spreadsheet itself is not documented fully, as there are no accompanying explanations on how each of the tabs function or how the Visual Basic processes produce the output. As the processing of the outputs is to some extent a "black box" the use of stress tests is necessary so that the expected results can be checked.

11.1 Implementation of formulae

The three types of elasticity implemented in the spreadsheet are as follows:

1. Constant (not dependent on the value of the variable);
2. Variable; and
3. Squared Terms.

The functional form that was assigned to each segment and geographical movement depends on the definition in the research. These corresponded with what was specified in the "I_Parameters" tab. Also in this tab are the lag parameter estimates from 'Revisting the Elasticity-Based Framework'. A small number of parameters were observed to be different between the implementation in the spreadsheet and the documentation. It is recommended that these parameters are checked further by DfT to ensure any errors are removed. Two examples are shown in Table 11.1:

Table 11.1: Lag Parameter discrepancies between spreadsheet and documentation

Segment	Parameter	Documentation	Spreadsheet
LSEE-LSEE full/reduced	Fare (two lags)	-0.00347	-0.0347
Other-Core full	Lag of Journey	0.259	0.268

For the implementation of the formulae, a cell-by-cell check was not undertaken, as the general forms of the calculations do not vary dramatically between segments with similar elasticity formulations. Instead, the calculation examples in the Guidance Document were recreated using the spreadsheet formulae for each of the three types of elasticity. The spreadsheet numbers matched the examples, giving reassurance that they were functioning in an identical fashion. The formulae in the spreadsheet were also cross checked against the algebra in the guidance document to eliminate the possibility that the correct results had been produced with the wrong calculations and/or numbers.

The calculation of the elasticity parameters, contained within the "C_Lags" tab, feed into the "C_Flow" tab, which is where the demand drivers and base data have the elasticities applied to produce forecasts. This tab uses a variety of two way lookups that are cycled through using Visual Basic, to change the Origin/Destination to produce the full output for all zone combinations (the "C_Flow" tab only illustrates a single zone pair at a time).

The "I_Sources" tab contains a comparison of the elasticities produced by the spreadsheet model versus those reported in the documentation. It can be seen that the constant elasticities are equal apart from slight rounding differences. The larger differences for the variable elasticities and squared term elasticities were a result of a discrepancy with the method of calculation. The difference is that the spreadsheet calculations do not multiply by the value of the segment, which accounts for their different orders of magnitude. As the segment values used by Oxera in the Findings Report (average value of segment in 2007/8) are not documented, it is not possible to compare the elasticities directly.

However, in order to spot check that the same variable elasticity results would be returned given the same value, a single known 2007/8 average segment value was used from the guidance document. This was applied using the calculations in the spreadsheet, which correctly returned the expected values. Calculation spot checks were undertaken for each of three types of elasticity, each of which returned values that were consistent between the Guidance Document and the spreadsheet formulae.

11.2 Stress Tests

Stress tests have been performed on the model to determine whether it is working as expected. The model was found to respond appropriately to changes in the following:

- elasticity parameter inputs;
- demand driver growth;
- elasticity type changes (constant, variable or squared); and
- the switching on and off of demand drivers.

The model was also checked to ensure that the growth index values (i.e. total growth in demand across all demand drivers from C_Flow tab) were correctly flowing through the workbook to the final output sheet (O_Summary). The 'input index values' are compared with the 'output index values' to determine if these are equal, which they should be for OD pairs with 'constant' elasticity parameters (e.g. Central London to Central Manchester). The model was found to correctly apply the growth index values.

Details and results of these tests are displayed in **Table 11.2**.

11.3 Summary

The functional form of each of the three types of elasticity was consistent with the Guidance document. The differences between the elasticities reported in the "I_Sources" tab related to rounding errors in the constant elasticities, and the absence of 2007/8 average values from the spreadsheet for direct comparison. This resulted in an inconsistent method of calculation that appears to be significant but is not the result of an incorrect implementation. The stress tests indicate that the model is responding appropriately to changes in inputs.

Table 11.2: Results of Stress Tests

				Central London to Central Manchester				
Test No.	**Description**	**Tab**	**Macro Run**	**Base Year Jnys**	**2010 Jnys**	**Output Index**	**Input Index**	**Check**
Test0	As received.	NA	✓	89,296	120,230	1.35	1.35	✓
Test1	Increased fare elasticity parameters by 10%.	I_Parameters	✓	89,296	120,230	1.35	1.35	✓
Test2	Set I_FAR demand driver growth to 5% from 1991 to 2010.	I_FAR	✓	89,296	286,733	3.21	3.21	✓
Test3	Set I_FAR demand driver growth to 6% from 1991 to 2010.	I_FAR	✓	89,296	239,520	2.68	2.68	✓
Test4	All demand drivers switched off.	I_Control	✓	89,296	89,296	1.00	1.00	✓
Test5	All demand drivers switched off except Population.	I_Control	✓	89,296	89,296	1.00	1.00	✓
Test6	Set I_CCO demand driver growth to 2% from 1991 to 2010.	I_CCO	✓	89,296	108,935	1.22	1.22	✓
Test7	Set I_CCO demand driver growth to 3% from 1991 to 2010.	I_CCO	✓	89,296	137,533	1.54	1.54	✓
Test8	Combination of Test2 and Test6.	I_FAR & I_CCO	✓	89,296	259,796	2.91	2.91	✓
Test9	Combination of Test3 and Test7.	I_FAR & I_CCO	✓	89,296	273,992	3.07	3.07	✓
Test10	All elasticity types set to 'constant'.	I_Parameters	✓	89,296	120,230	1.35	1.35	✓
Test11	As Test10 but fare elasticity type for 'Other-LSEE_S' changed to 'variable'.	I_Parameters	✓	89,296	120,230	1.35	1.35	✓

12 Ticket Aggregation and Car Journey Times

12.1 Car journey time data

There were two queries raised about the car journey time measure, which have been addressed in this update phase to the Revisiting study:

- the car journey time variable used in the econometric analysis was a logged growth rate; and
- the car cost variable varies by journey purpose, while the car journey time variable does not.

The use of logged growth rates in the analysis was unintended. However, analysis using the correct variable, i.e., logged levels of car journey time, suggests that they were not statistically or economically significant and so this issue probably had relatively little impact on the results of the initial study.

That the car cost variable was disaggregated by journey purpose, while car journey time was not, was not an error, but the underlying rationale merits further explanation. The car journey time variable used estimates of the number of trips by TEMPRO zone, taken from NTEM v5.4, combined with a speed-flow curve to estimate the growth in car journey time, split by flow and year. These growth rates were applied to base journey time data from the National Transport Model (NTM) to obtain journey times. This process is described in the Appendix to the Data Capability report.

The car journey time growth rates (which are not disaggregated by time period) were based on morning peak speed-flow curves, in order to provide growth rates in car journey time which were appropriate, in particular, for business and commuting. It was judged that the benefits of having growth rates split by journey purpose or time of day were not sufficient to justify the costs of creating the new variables, given the computational demands and the consequent delay to the remainder of the analysis.

GIS techniques were then used to derive the distance travelled through each TEMPRO zone which, when combined with the car journey time levels, provided the data required to calculate journey speed. Speed-cost curves from the NTM and Transport Model for Scotland (split by journey purpose) were then used to generate the car cost variable.

12.2 Ticket-type aggregation

In the initial Revisiting study, a single model was estimated for full and reduced tickets within the LSEE to LSEE market segment. The rationale for this combination of ticket types was that the history of the segment is such that there was a considerable degree of switching between full and reduced price tickets for reasons which were not related to price. The DfT requested that the way in which the full and reduced price ticket yields were aggregated was changed from the approach adopted in the original analysis, and then for the econometric analysis to be rerun on the revised data. The results of this are in **Table 12.1**, below. More details and a full dashboard of the original and revised models are contained in Appendix A.

Table 12.1: LSEE-to-LSEE, combined full and reduced

	One-year elasticities		Three-year elasticities	
	Revised	**Original**	**Revised**	**Original**
Fare	**–0.48**	–0.79	**–0.56**	–0.95
Income	**0.14**	–0.74	1.72	1.58
Employment	0.46	0.48	0.49	0.49
Car cost	**0.43**	0.82	**0.92**	1.44
GJT	–0.36	–0.39	–1.52	–1.60
Performance	**0.91**	0.43	**0.75**	1.14

Note: GJT stands for Generalised Journey Time. Numbers in bold lie outside the 95% confidence interval of the original results.

Source: Oxera analysis.

As can be seen from the table above, the change in the way in which the yield is aggregated has a substantial effect on a number of the elasticities. The general impact of this change is to reduce (in absolute magnitude) the elasticities of rail demand to fare, car cost and performance after three years. The elasticities of income, employment and GJT remain broadly unchanged.

Appendix A

Dashboards for Revised Analysis

15 February 2012

Definitions

Segments

London, South East and East of England (LSEE)	London, South East and the East of England are defined by the relevant Government Office Regions (GORs).
Core cities	Core cities are defined as non-London core cities. Specifically: Birmingham, Manchester, Liverpool, Nottingham, Bristol, Sheffield, Cardiff, Edinburgh, Glasgow, Newcastle, Leeds, Leicester, York and Hull.
Airports	The airport stations that are included within the airports segment are: Gatwick Airport, Stansted Airport, Luton Airport Parkway, Manchester Airport, Birmingham International, Inverkeithing (Edinburgh) and Prestwick International Airport.
Other	The other segment includes all the flows in the dataset that are not included in one of the other segments.

Summary statistics

Market share (journeys)	Market share is calculated as journeys within the segment as a proportion of the total journeys in the dataset.
Market share (distance)	Market share is calculated as passenger journeys multiplied by the average distance of the flows within the segment as a proportion of the total dataset.
Market share (revenue)	Market share is calculated as revenue within the segment as a proportion of the total revenue in the dataset.
Passenger journeys	Average annual number of journeys for the segment. Note that this is affected by lower journeys at the beginning of the sample period.
Average distance	Average distance is measured in km.
Average fare per km	Average fare per km is calculated as the average fare divided by the average distance within the segment. Units are pence per km (2007 prices).

Diagnostics

Model formulation	The estimator used in this project is the Blundell and Bond estimator for dynamic panel data, which is a system Generalised Method of Moments (GMM) estimator for dynamic panel data, for which there are no direct tests of model mis-specification. However, there are a number of diagnostic tests which can be carried out to indirectly test for model mis-specification. These are set out below.
Arellano–Bond (autocorrelation)	This tests for autocorrelation in the first differenced error terms of the regression.
Sargan (instrument validity)	This tests the validity of the underlying assumptions of the estimator. However, it can still be rejected if there is heterogeneity in the data-generating process, even if the model is correctly specified.
Unit root test on residuals	If the error term from a model is non-stationary, then the identified relationships which have been identified between the variables may be spurious. To test this, the panel data unit root test developed by Maddala and Wu (1999) has been used.
Stable model	The stability of the models has been tested by rolling regressions.
Model fit	Unlike in ordinary least squares (OLS) regression, there is no R2 for system GMM estimators. As a measure of model fit, the squared correlation between actual and fitted data is presented. This measure is bounded by zero and one, with a measure of one showing perfect correlation between actual and fitted data.

Core cities to London, South East and East of England: Reduced price (revised)

Variable list	One-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after one year	Three-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after three years
Fare	**-0.845**	**-1.244**	**-0.446**	**79%**	**-1.057**	**-1.553**	**-0.562**	**99%**
Cross-price								
Income	**-1.196**	**-2.014**	**-0.378**	**-47%**	**2.393**	**1.172**	**3.614**	**94%**
Population								
Employment								
Car ownership								
Car cost	**0.496**	**0.229**	**0.764**	**79%**	**0.621**	**0.288**	**0.953**	**99%**
Car Journey Time								
GJT	*-0.444*	-1.026	0.138	79%	*-0.555*	-1.269	0.158	99%
Performance	0.541	0.213	0.869	79%	0 677	0.278	1.076	99%
SQI								

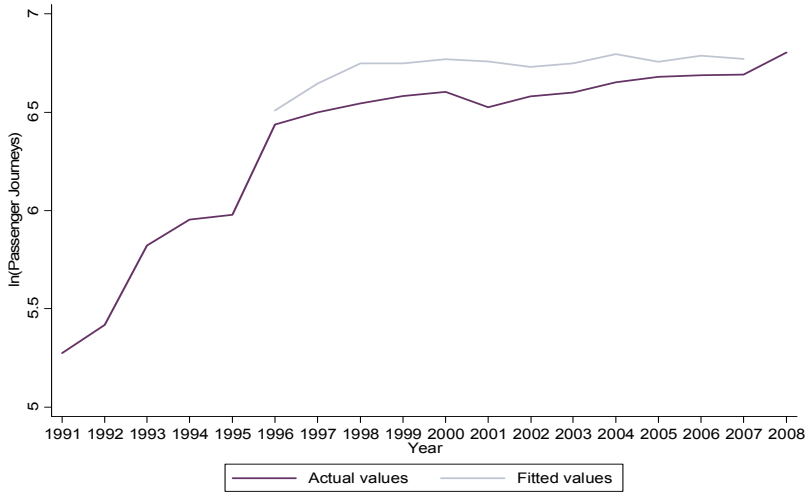
*The variables in bold are the elasticities which are of direct interest for the project

Summary statistics for segment

Market share (Journeys)	0.7%
Market share (Distance)	4.5%
Market share (Revenue)	2.9%
Passenger journeys	4,356,000
Ave distance (km)	189.2
Ave fare per km (£)	0.113

Diagnostics

Model formulation	Constant elasticities
Sample size (number of observations)	7143
Number of years of sample	12
Arellano–Bond (autocorrelation)	Pass
Sargan (instrument validity)	Fail
Unit root test on residuals	Could not conduct unit root test: too many values
Stable model	Pass
Model fit	0.65



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys) over time

Comments on diagnostics: The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on model

Variable definition

Fare	Revenue/journeys
Cross-price	
Income	Disposable income per capita at origin
Population	
Employment	
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	Sectoral PPM
SQI	

Core cities to London, South East and East of England: Reduced price (original)

Variable list	One-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after one year	Three-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after three years
Fare	**-0.56**	**-0.881**	**-0.232**	**81%**	**-0.68**	**-1.084**	**-0.285**	**99%**
Cross-price								
Income	***-0.55***	**-1.498**	**0.406**	**-37%**	**1.40**	**0.488**	**2.311**	**95%**
Population								
Employment								
Car ownership								
Car cost	**0.43**	**0.100**	**0.752**	**81%**	**0.52**	**0.122**	**0.928**	**99%**
Car Journey Time								
GJT	-0.94	-1.456	-0.416	81%	-1.15	-1.762	-0.543	99%
Performance	*0.31*	-0.0223	0 645	81%	*0.38*	-0.0268	0.794	99%
SQI								

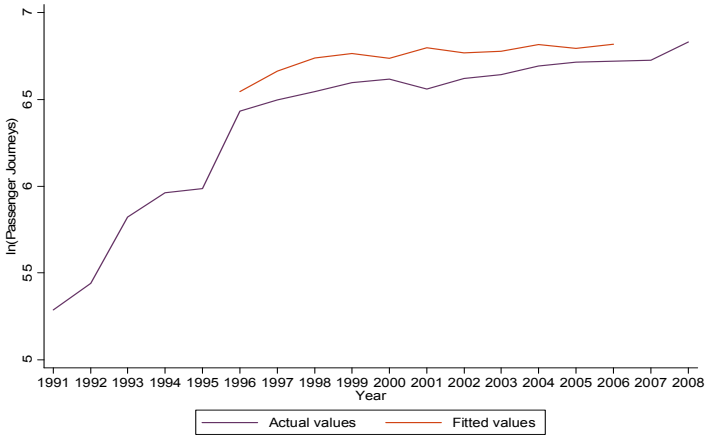
*The variables in bold are the elasticities which are of direct interest for the project

Summary statistics for segment

Market share (Journeys)	0.7%
Market share (Distance)	4.5%
Market share (Revenue)	2.9%
Passenger journeys	4,356,000
Ave distance (km)	189.2
Ave fare per km (£)	0.113

Diagnostics

Model formulation	Constant elasticities
Sample size (number of observations)	6868
Number of years of sample	11
Arellano–Bond (autocorrelation)	Pass
Sargan (instrument validity)	Fail
Unit root test on residuals	Pass
Stable model	Pass
Model fit	0.64



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys) over time

Comments on diagnostics: The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on model

Variable definition

Fare	Revenue/journeys
Cross-price	
Income	Disposable income per capita at origin
Population	
Employment	
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	Sectoral PPM
SQI	

Other to London, South East and East of England: Full price (revised)

Variable list	One-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after one year	Three-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after three years
Fare	**-1.814**	**-2.055**	**-1.572**	**123%**	**-1.501**	**-1.818**	**-1.183**	**102%**
Cross-price	**0.375**	**0.215**	**0.535**	**73%**	**0.505**	**0.296**	**0.714**	**98%**
Income	**0.673**	**0.0952**	**1.250**	**73%**	**0.907**	**0.153**	**1.661**	**98%**
Population	**1.880**	**0.288**	**3.472**	**73%**	**2.535**	**0.409**	**4.662**	**98%**
Employment								
Car ownership								
Car cost	**0.693**	**0.238**	**1.148**	**73%**	**0.935**	**0.315**	**1.554**	**98%**
Car Journey Time								
GJT	-0.154	-0.440	0.132	12%	-1.168	-1.699	-0.637	93%
Performance								
SQI	1.161	0.689	1.633	73%	1.565	0.910	2.220	98%

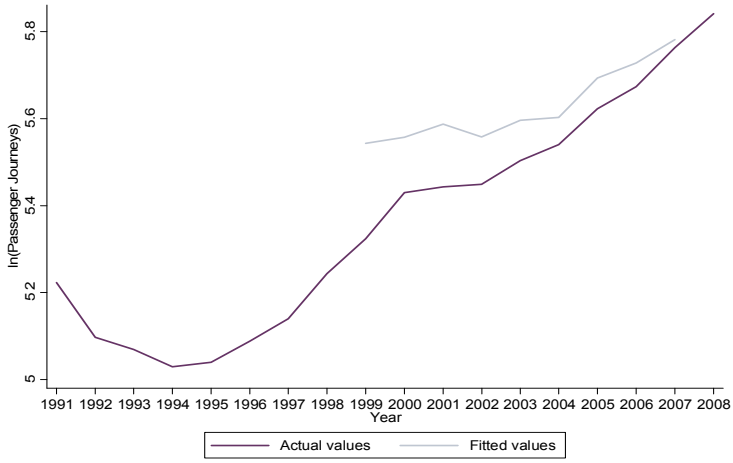
*The variables in bold are the elasticities which are of direct interest for the project

Summary statistics for segment

Market share (Journeys)	0.6%
Market share (Distance)	2.9%
Market share (Revenue)	5.9%
Passenger journeys	3,518,000
Ave distance (km)	166.7
Ave fare per km (£)	0.272

Diagnostics

Model formulation	Constant elasticities
Sample size (number of observations)	13937
Number of years of sample	9
Arellano–Bond (autocorrelation)	Pass
Sargan (instrument validity)	Fail
Unit root test on residuals	Pass
Stable model	Pass
Model fit	0.55



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys) over time

Comments on diagnostics: The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on model

Variable definition

Fare	Revenue/journeys
Cross-price	Reduced price tickets
Income	Disposable income per capita at origin
Population	Total population at origin
Employment	
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	
SQI	Service Quality Index

Other to London, South East and East of England: Full price (original)

Variable list	One-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after one year	Three-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after three years
Fare	**-1.81**	**-2.055**	**-1.572**	**123%**	**-1.50**	**-1.817**	**-1.180**	**102%**
Cross-price	**0.38**	**0.216**	**0.535**	**73%**	**0.51**	**0.298**	**0.714**	**98%**
Income	**0.69**	**0.114**	**1.269**	**73%**	**0.93**	**0.180**	**1.685**	**98%**
Population	**1.88**	**0.281**	**3.469**	**73%**	**2.53**	**0.398**	**4.657**	**98%**
Employment								
Car ownership								
Car cost	**0.68**	**0.227**	**1.131**	**73%**	**0.91**	**0.300**	**1.531**	**98%**
Car Journey Time								
GJT	-0.16	-0.446	0.125	13%	-1.18	-1.712	-0.654	93%
Performance								
SQI	1.15	0.678	1.620	73%	1 55	0.895	2.203	98%

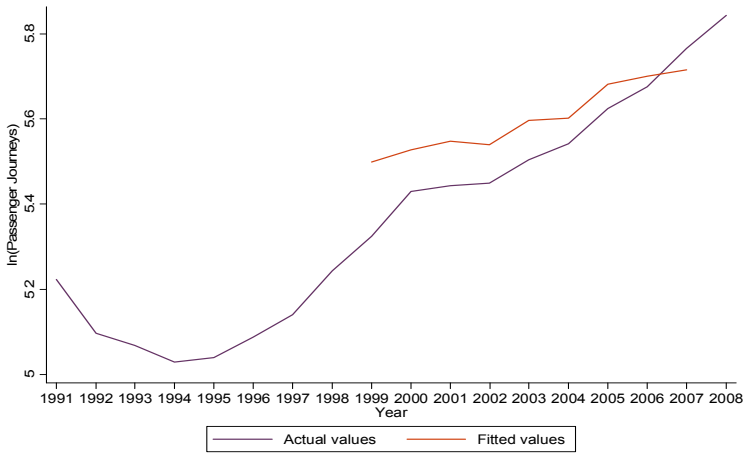
*The variables in bold are the elasticities which are of direct interest for the project

Summary statistics for segment

Market share (Journeys)	0.6%
Market share (Distance)	2.9%
Market share (Revenue)	5.9%
Passenger journeys	3,518,000
Ave distance (km)	166.7
Ave fare per km (£)	0.272

Diagnostics

Model formulation	Constant elasticities
Sample size (number of observations)	13919
Number of years of sample	9
Arellano–Bond (autocorrelation)	Pass
Sargan (instrument validity)	Fail
Unit root test on residuals	Pass
Stable model	Pass
Model fit	0.55



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys) over time

Comments on diagnostics: The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on model

Variable definition

Fare	Revenue/journeys
Cross-price	Reduced price tickets
Income	Disposable income per capita at origin
Population	Total population at origin
Employment	
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	
SQI	Service Quality Index

Other to Other: Season tickets (revised)

Variable list	One-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after one year	Three-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after three years
Fare	**-0.970**	**-1.172**	**-0.768**	**37%**	**-1.328**	**-1.906**	**-0.749**	**50%**
Cross-price								
Income	***-0.502***	**-1.234**	**0.231**	**11%**	***-1.340***	**-3.302**	**0.622**	**30%**
Population								
Employment								
Car ownership								
Car cost	***-0.424***	**-0.852**	**0.00325**	**345%**	***-0.359***	**-1.189**	**0.471**	**292%**
Car Journey Time								
GJT	*-0.0433*	-0.255	0.169	1%	*-0.672*	-1.515	0.171	23%
Performance	0.780	0.248	1.312	5%	3.974	2.102	5.845	24%
SQI								

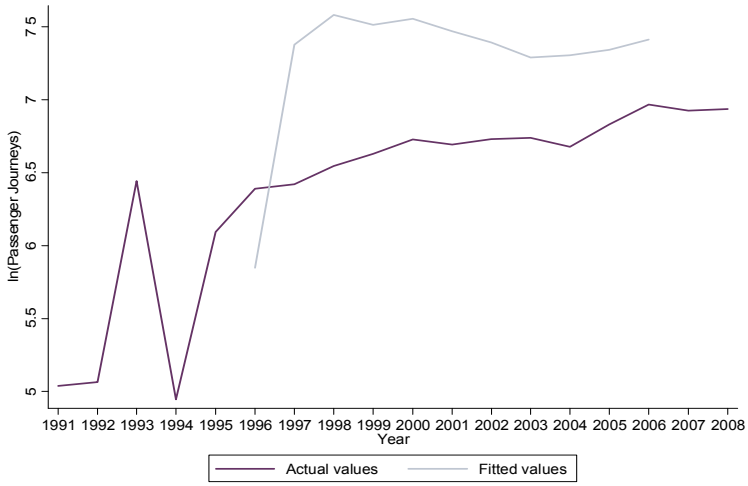
*The variables in bold are the elasticities which are of direct interest for the project

Summary statistics for segment

Market share (Journeys)	1.0%
Market share (Distance)	0.6%
Market share (Revenue)	0.3%
Passenger journeys	5,321,000
Ave distance (km)	45 27
Ave fare per km (£)	0.114

Diagnostics

Model formulation	Constant elasticities
Sample size (number of observations)	10,539
Number of years of sample	12
Arellano–Bond (autocorrelation)	Pass
Sargan (instrument validity)	Fail
Unit root test on residuals	Cannot conduct unit root test
Stable model	Pass
Model fit	0.79



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys) over time

Comments on diagnostics: The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on model

Variable definition

Fare	Revenue/journeys
Cross-price	
Income	GVA per employee at destination
Population	
Employment	
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	Sectoral PPM
SQI	

Other to Other: Season tickets (original)

Variable list	One-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after one year	Three-year elasticity: Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after three years
Fare	**-0.99**	**-1.192**	**-0.785**	**44%**	**-1.32**	**-1.868**	**-0.769**	**59%**
Cross-price								
Income	***-0.58***	**-1.308**	**0.152**	**14%**	***-1.50***	**-3.382**	**0.390**	**37%**
Population								
Employment								
Car ownership								
Car cost	***-0.82***	**-1.344**	**-0.286**	**-559%**	***-0.56***	**-1.483**	**0.367**	**-383%**
Car Journey Time								
GJT	*-0.11*	-0.363	0.143	3%	*-0.97*	-2.015	0 0712	29%
Performance	-0.23	-0.587	0.133	-2%	2 68	1.364	4	18%
SQI								

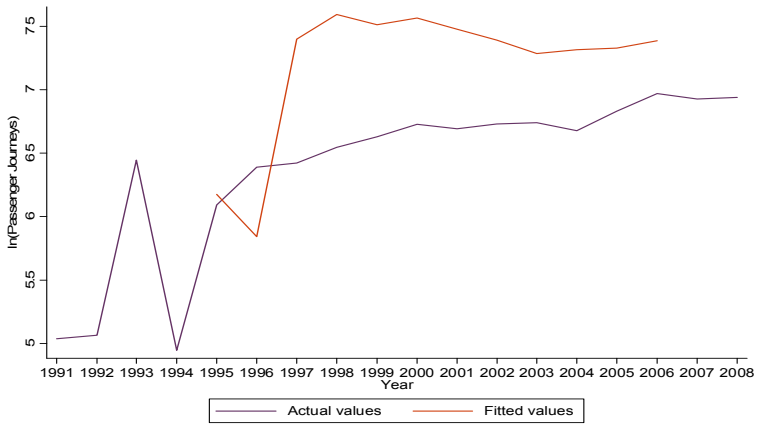
*The variables in bold are the elasticities which are of direct interest for the project

Summary statistics for segment

Market share (Journeys)	1.0%
Market share (Distance)	0.6%
Market share (Revenue)	0.3%
Passenger journeys	5,321,000
Ave distance (km)	45 27
Ave fare per km (£)	0.114

Diagnostics

Model formulation	Constant elasticities
Sample size (number of observations)	10,590
Number of years of sample	12
Arellano–Bond (autocorrelation)	Pass
Sargan (instrument validity)	Fail
Unit root test on residuals	Could not conduct unit root test
Stable model	Pass
Model fit	0.76



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys) over time

Comments on diagnostics: The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on model

Variable definition

Fare	Revenue/journeys
Cross-price	
Income	GVA per employee at destination
Population	
Employment	
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	Sectoral PPM
SQI	

London, South East and East of England to London, South East and East of England: Combined model (revised)

Variable list	One-year elasticity: Point estimate	One-year elasticity: 95% CI lower bound	One-year elasticity: 95% CI upper bound	One-year elasticity: Percentage adjustment to long run after one year	Three-year elasticity: Point estimate	Three-year elasticity: 95% CI lower bound	Three-year elasticity: 95% CI upper bound	Three-year elasticity: Percentage adjustment to long run after three years
Fare	**-0.477**	**-0.694**	**-0.259**	**84%**	**-0.562**	**-0.745**	**-0.380**	**99%**
Cross-price								
Income	**0.137**	**-0.182**	**0.455**	**8%**	**1.719**	**1.429**	**2.008**	**100%**
Population								
Employment	**0.464**	**0.363**	**0.565**	**95%**	**0.487**	**0.383**	**0.590**	**100%**
Car ownership								
Car cost	**0.429**	**0.332**	**0.527**	**46%**	**0.921**	**0.762**	**1.080**	**99%**
Car Journey Time								
GJT	-0.357	-0.467	-0 248	23%	-1 524	-1.714	-1.333	98%
Performance	0.906	0.765	1 047	120%	0.754	0 599	0 909	100%
SQI								

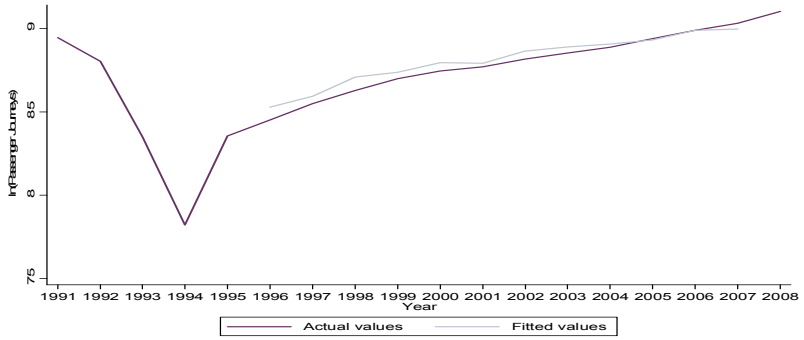
are the elasticities which

Summary

Market share	34.9%
Market share	24.1%
Market share	26.5%
Passenger	193,000,000
Ave distance (km)	30.1
Ave fare per km (£)	0.194

Diagnos ics

Model formulation	Constant
Sample size	78496
Number of years of	12
	0
Arellano–Bond	Pass
Sargan (instrument	Fail
Unit root test on	Could not conduct
Stable model	Pass
Model fit	0.47



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys)

Comments on The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on

Variable

Fare	Revenue/journeys
Cross-price	
Income	Disposable income per capita at origin
Population	
Employment	Total jobs at destination
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	Sectoral PPM
SQI	

London, South East and East of England to London, South East and East of England: Combined model

Variable list	One-year elasticity				Three-year elasticity			
	Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after one year	Point estimate	95% CI lower bound	95% CI upper bound	Percentage adjustment to long run after three years
Fare	**-0.79**	**-0.899**	**-0.675**	**83%**	**-0.95**	**-1.070**	**-0.834**	**100%**
Cross-price								
Income	**0.74**	**0.425**	**1.059**	**47%**	**1.58**	**1.382**	**1.787**	**100%**
Population								
Employment	**0.48**	**0.369**	**0.584**	**97%**	**0.49**	**0.384**	**0.601**	**100%**
Car ownership								
Car cost	**0.82**	**0.726**	**0.909**	**57%**	**1.44**	**1.291**	**1.579**	**99%**
Car Journey Time								
GJT	-0 39	-0 501	-0 286	24%	-1.60	-1 804	-1.402	98%
Performance	0.43	0.336	0 514	37%	1.14	1 040	1 246	100%
SQI								

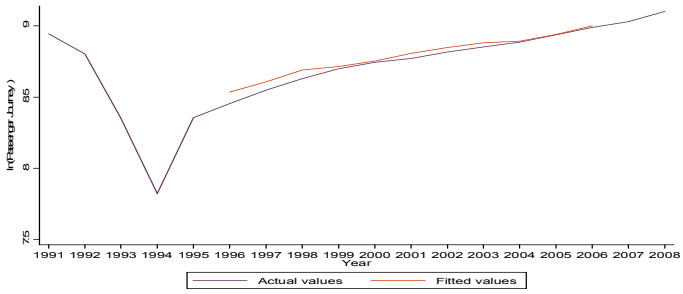
are the elasticities which

Summary

Market share	34.9%
Market share	24.1%
Market share	26.5%
Passenger	193,000,000
Ave distance (km)	30.1
Ave fare per km (£)	0.211

Diagnos ics

Model formulation	Constant
Sample size	71894
Number of years of	11
Arellano–Bond	Pass
Sargan (instrument	Fail
Unit root test on	Could not conduct
Stable model	Pass
Model fit	0.42



As a measure of goodness of fit, the graph illustrates how well the predicted values from the model match the actual values of log(journeys)

Comments on The Sargan test may still be failed if there is heterogeneity in the DGP, even if the model specification is correct

Comments on

Variable

Fare	Revenue/journeys
Cross-price	
Income	Disposable income per capita at origin
Population	
Employment	Total jobs at destination
Car ownership	
Car cost	Cost of journey
Car Journey Time	
GJT	Generalised Journey Time
Performance	Sectoral PPM
SQI	