RESEARCH ON LONG TERM FARE ELASTICITIES

A Final Report for the Strategic Rail Authority

Prepared by NERA

November 2003
London

Project Team:
Ian Jones
Jan Peter van der Veer
Daniel Paredes
Paul Metcalfe
Prof. John Cubbin
Dr. Dimitrios Asteriou

n/e/r/a

National Economic Research Associates
Economic Consultants

15 Stratford Place
London W1C 1BE
Tel: (+44) 20 7659 8500
Fax: (+44) 20 7659 8501
Web: http://www.nera.com

An MMC Company

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Background to the Research

This report, prepared by National Economic Research Associates (NERA) in conjunction with Professor John Cubbin and Dr Dimitrios Asteriou of City University, presents estimates of the long term elasticities of demand for passenger rail travel in Great Britain. The background to the project is that there is a lack of clarity regarding the timescale over which the current elasticities in the PDFH are intended to apply. For certain uses of the PDFH, such as investment appraisal and long-term planning, it would be desirable to have long-term elasticities. In other contexts, notably short term business and financial planning, rail business managers require advice on the response of demand to price changes in the short run, defined as a period of about a year after the fare change has occurred. The PDFH indicates that the elasticity values it recommends are intended to represent the change in demand that occurs within a year of a change in fares. This definition implies that demand effects occurring in the longer term are not addressed. It also implies that the elasticities do not offer guidance on the short run response of demand within the year.

In the present project, we have undertaken extensive econometric analysis on large data sets which has enabled us not only to derive long-term fare elasticities, but also to examine the short run response of demand to changes in price and other factors.

The current PDFH elasticities are grouped by market segment. For a number of market segments, the recommended values are highly disaggregated, for example by ticket type category, length of journey and fare per mile paid. For other market segments, the PDFH elasticity values are of a more aggregate nature.

Previous Studies

There have been a number of studies of long term demand elasticities in various sectors of the economy, including rail travel. Studies on long-term rail demand elasticities have been undertaken by Owen and Phillips (1987) and NERA (1999). Recently, work has been undertaken by Dargay and Hanly (2002) on long-run elasticity values of the demand for local bus travel. There have also been studies on long-run elasticity values of petrol demand, some of which, like the present study, have used cointegration techniques that allow analysis of short term adjustment of demand as well as long term elasticities to estimate the speed of adjustment in demand.

Trends in Demand and Revenue

Before undertaking our own econometric estimation work, we have reviewed previous trends in demand and average revenue in each of the main market segments. This review produced some unexpected insights. For example, in the long-distance market, there has been a spectacular growth in the number of “other” tickets, mainly consisting of operator-specific tickets introduced after privatisation. This development has had important implications for our estimation work. Also, we have established that the fall in journeys on long-distance services during the early 1990s has been entirely due to reductions in short to medium distance commuting to London on long-distance services. The amount of long-distance ***travel*** during this period was relatively stable.

Data

We have collected data from a number of sources. The most important source has been the CAPRI database, from which we have obtained journey and revenue data for the flows that we have considered. From other sources, we have collected data on GDP; generalised journey time; service quality (punctuality and reliability); and attributes of other modes. We have also constructed an ***economic activity*** variable, representing the level of cyclical activity in the economy.

Market Segmentation

We have carried out our analysis for four distinct panels, based on the market segments in the PDFH. The panels are:

- London TCA to and from the Rest of Country (other than the South East);
- Non London Long Distance;
- London TCA and South East; and
- Non London Short Distance.

For each panel, we have carried out analyses both at the aggregate and at the disaggregate level. For the two long distance panels, we have distinguished between fully flexible (First and Full Standard) and reduced/restricted tickets. Our analysis of the two short distance panels distinguishes between season and non-season tickets. Due to potential multicollinearity problems in our dataset, all our disaggregated results need to be interpreted with caution.

Econometric Methodology

The choice of econometric model has been driven by the fact that we have established a ***cointegrating relationship*** between the variables in our model. Moreover, since we were concerned with estimating elasticities for whole panels rather than for individual flows, we chose a ***panel fixed effects*** framework, which constrains the elasticity parameters to be the same for each flow within panels. In addition to the panel fixed effects models, we have also obtained panel mean group estimates for ***Vector Error Correction Models*** (VECMs) for each panel. The VECM approach has enabled us to examine the speed with which demand adjusts towards the long run equilibrium level following a change in fares.

For each panel, we have tested for the significance of the explanatory variables for which we had collected data. Our measures of generalised journey time, service quality and the attributes of alternative transport modes were each found to be insignificant. The economic activity was significant only for the London TCA and South East panel and has been included in the equations for this panel. The other core explanatory variables are price, GDP, seasonal effects dummy variables and two dummy variables capturing the impact of the Hatfield accident.

Results

The key results of our analyses of the two long distance panels are summarised in Table 1 below. The table contains the following:

- aggregate long-term elasticity estimates for each panel from our panel fixed effects model;
- long-term elasticity estimates from our panel fixed effects model disaggregated by ticket type category (fully flexible and reduced / restricted); and
- short (covering the demand response within four weeks of the fare change) and long term aggregate elasticity estimates from the VECM analysis, as well as the proportion of the change in demand occurring within one year.

Table 1
NERA Elasticity Estimates: Long Distance

	Panel fixed effects (long term)			VECM (short and long term; aggregate only)		
	Aggregate	Fully flexible	Reduced/ restricted	***Short term***	Long term	Proportion of demand change occurring within one year
London TCA to and from Rest of Country	-0.64	-0.49	-0.82	***-0.20***	-0.70	99%
Non London Long Distance	-0.82	-1.16	-0.70	***-0.18***	-1.01	92%

Source: NERA estimates

Compared to the PDFH recommended values (which are highly disaggregated), the NERA estimates for the London TCA to and from Rest of Country panel are considerably lower. For the Non London Long Distance panel, the NERA estimates are of the same order of magnitude as the PDFH elasticities.

Our key estimates for the two short-distance panels are summarised in Table 2.

Table 2
NERA Elasticity Estimates: Short Distance

	Panel fixed effects (long term)			VECM (short and long term; aggregate only)		
	Aggregate	Seasons	Non-seasons	***Short term***	Long term	Proportion of demand change occurring within one year
London TCA and South East	-0.62	-0.83	-0.40	***-0.16***	-0.68	97%
Non-London Short Distance	-0.95	-	-	***-0.34***	-1.14	99%

Source: NERA estimates

The NERA estimates for these two panels cannot easily be compared with the PDFH recommendations in view of differences in the extent of disaggregation between the two. However, in the London TCA to and from Rest of Country panel, the estimated season ticket elasticity is considerably higher than the recommendation in the PDFH, particularly for traffic to London. In the Non London Short Distance panel, the PDFH recommendations are also generally below our estimated long-term values.

Implications for PDFH Elasticities

Given that demand adjusts so quickly, the PDFH elasticities can be regarded as being very close to long run elasticities. However, even under that assumption, our estimated elasticity values are as noted above in some cases different from the PDFH recommendations. Since we have not had access to the studies underlying the PDFH, it is not always possible to comment on these differences.

The key implications of our research for the PDFH recommended elasticity values are in our view the following:

- In the London TCA to and from Rest of Country panel, our estimated aggregate elasticity is well below the elasticities in the PDFH. In our 1999 study for OPRAF, we found a similar result. In view of the results of our work, we would recommend a short comparison study of our work with the study underlying the PDFH values. Based on that, a decision could be made as to whether the PDFH values should be adjusted downwards.

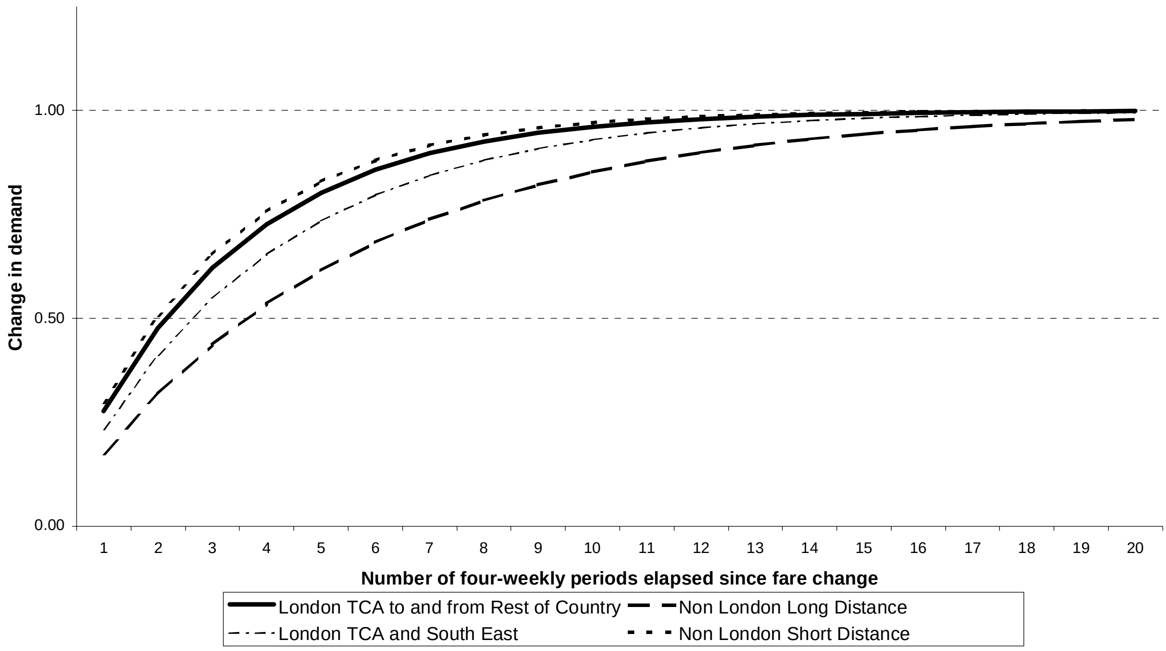
- Our results suggest substantial differences between ticket type categories in the Non London Long Distance panel, where the PDFH does not currently provide disaggregated recommendations. While we are not sufficiently confident in our disaggregated results to propose them as a basis for PDFH values, our results do suggest a possible need for further research in this area.

- In the London TCA and South East panel, there are significant differences between the PDFH recommended values and our disaggregated results. Whilst we would not argue that our disaggregated results should form the basis for PDFH recommendations, we believe that they do justify a more detailed re-appraisal of the PDFH parameter values.

- In a number of panels, particularly those where elasticity values are differentiated according to the fare per mile paid, we believe there is a case for simplifying the recommendations.

If the PDFH elasticities are interpreted as long-run elasticities, then the results on the speed of adjustment of demand may be used to provide guidance in the PDFH on short run demand responses. Such guidance could be used as input to short term planning and budgeting procedures. The speed of adjustment of demand in each of our panels is shown graphically in Figure 1 below.

Figure 1
Speed of Adjustment in Demand by Market Segment
(assuming the long-term change in demand is 1 per cent)



Source: NERA estimates

Finally, our report contains recommendations for the development of an updating mechanism. In our view, the updating mechanism should consist of a three-stage process:

- Stage 1: Reviewing the purposes for which elasticity estimates are required. Are the current estimates appropriate in principle to the demands likely to be made on them?

- Stage 2: If they remain appropriate, deciding whether at any point in time an update of the elasticities should in fact take place.
- Stage 3: The derivation of the new elasticity values.

The report provides detailed recommendations for the procedures to be followed in each of these stages.

1. INTRODUCTION

This report, prepared by National Economic Research Associates (NERA) in conjunction with Professor John Cubbin and Dr Dimitrios Asteriou of City University, presents estimates of the long term elasticities of demand for passenger rail travel in Great Britain. The background to the project is that there is a lack of clarity regarding the timescale over which the current elasticities in the PDFH are intended to apply. For certain uses of the PDFH, such as investment appraisal and long-term planning, it would be desirable to have long-term elasticities. In other contexts, notably short term business and financial planning, rail business managers require advice on the response of demand to price changes in the short run, defined as a period of about a year after the fare change has occurred.

In the present project, we have undertaken extensive econometric analysis on large data sets which has enabled us not only to derive long-term fare elasticities, but also to examine the short run response of demand to changes in price and other factors.

Our analysis has used data on revenues and journey volumes extracted from the CAPRI database, supplemented with data from other sources. While the CAPRI database is an invaluable source of information, it needs to be used extremely carefully. We have found evidence of serious recording errors, in particular during periods following major disruptions to rail services, such as occurred following the Hatfield accident. Changes in ticket restrictions, the introduction and withdrawal of day returns and similar changes can limit the scope for disaggregate analysis. Finally, the development of novel types of ticket by individual TOCs since privatisation, means that the database has become far more complex, with increased scope for recording errors. These changes also have the effect of reducing the “visibility” of the data, so that it is no longer possible to link many of the new ticket types to existing ticket categories. This again restricts the feasibility of carrying out disaggregated analysis.

We have spent a considerable amount of time examining, verifying and where necessary cleaning the data. Nevertheless, the results of any disaggregated time-series analysis on the basis of CAPRI data must be interpreted with caution, including our own results.

The structure of the report is as follows:

- Chapter 2 presents a survey of earlier work on demand elasticities. It discusses the studies underlying the current PDFH recommended values. It also describes the findings of earlier studies that have reported estimates of long term demand elasticities, including studies of the demand for passenger rail services, and for other types of transport.
- Chapter 3 examines recent trends in rail traffic and in some of the variables that are likely to affect the demand for rail travel.

- Chapter 4 describes the process of assembling the data to be used in the econometric analysis, and sets out the specification of the econometric models that we have used to derive estimates of long term demand elasticities. It also discusses the approach to market segmentation used in our analysis, which corresponds to the market segmentation applied in PDFH.
- Chapter 5 presents and discusses the results of our econometric analysis, including the findings on both long term elasticities, and short run demand responses for each market segment.
- In Chapter 6, we discuss the status of the current PDFH recommended elasticity values. We discuss the nature of the PDFH elasticities in the light of the findings from our dynamic demand modelling. We then consider the implications of our estimates of long term elasticities for the PDFH values. We also show how the results of our dynamic modelling of rail demand can be used to provide guidance on short term demand responses.
- Chapter 7 contains our recommendations for the development of an updating mechanism.

The report contains the following appendices:

- Appendix A gives details on the flows contained in each panel;
- Appendix B discusses the tests for cointegration that we have applied;
- Appendix C provides detailed estimation results for each model that we have examined;
- Appendix D contains an analysis of the extent of "multicollinearity" in the data;
- Appendix E contains an example of the prediction tests that are available in the EViews software package; and
- Appendix F lists the references that we have used.

Our thanks are due to the various people within SRA and the TOCs that sometimes spent considerable amounts of time assisting us with obtaining and verifying the data.

2. SURVEY OF EARLIER WORK

2.1. Studies Underlying the Current PDFH Recommendations

2.1.1. Introduction

The current PDFH recommended elasticities are largely based on two studies. These are:

- The Rail Industry Forecasting Framework (RIFF), undertaken by Steer Davies Gleave and completed in December 1999.
- The Handbook Update Demand Analysis study by ITS Leeds, completed in May 2002.

The studies are summarised in PDFH, but the SRA has told us that no detailed project reports are available for either study. However, we have received some supplementary material from the researchers involved in response to our requests for information.

2.1.2. London TCA to and from Rest of Country

Elasticities for this market segment are disaggregated by ticket type category and are shown in Table 2.1. The Table distinguishes between various types of elasticities:-

- So-called total fares elasticities, f_t ,shown in the first column. These total elasticities represent the elasticity of demand aggregated across all ticket types with respect to an across-the-board change in ticket prices. The total fare elasticities are derived as the ticket volume-weighted average of the so-called conditional elasticities shown in the final column of the table, which represent the own price elasticities of demand for each ticket type observed if the prices of all ticket types are changed by the same proportional amount.
- Own price elasticities of demand, $f_{11,}$ $f_{FF,\ e}$ which would apply if the prices of all other tickets remained unchanged, and shown as the diagonal elements in Table 2.1.
- Cross price elasticities of demand, shown by the off-diagonal elements, $f_{F1,}$ in Table 2.1, representing the response of demand for a particular type of ticket to a change in price of another ticket type.

Table 2.1
Recommended Elasticity Values for London TCA to and from Rest of Country

	Total		1st		Full		Red		Cond
	20 to 100 miles								
Total	f_t	-0.9	-	-	-	-	-	-	-
1st	-	-	f_{11}	-0.80	f_{1F}	0.30	F_{1R}	0.00	-0.50
Full	-	-	f_{F1}	0.25	f_{FF}	-1.20	F_{FR}	0.35	-0.60
Red	-	-	f_{R1}	0.00	f_{RF}	0.25	F_{RR}	-1.25	-1.00
	101-200 miles								
Total	f_t	-1.0	-	-	-	-	-	-	-
1st	-	-	f_{11}	-0.90	f_{1F}	0.40	F_{1R}	0.00	-0.50
Full	-	-	f_{F1}	0.45	f_{FF}	-1.50	F_{FR}	0.45	-0.60
Red	-	-	f_{R1}	0.00	f_{RF}	0.30	f_{RR}	-1.40	-1.10
	Over 200 miles								
Total	f_t	-1.0	-	-	-	-	-	-	-
1st	-	-	F_{11}	-0.60	F_{1F}	0.10	f_{1R}	0.00	-0.50
Full	-	-	F_{F1}	0.15	F_{FF}	-1.00	f_{FR}	0.25	-0.60
Red	-	-	F_{R1}	0.00	F_{RF}	0.10	f_{RR}	-1.25	-1.15

These elasticities have been estimated in the Handbook update demand analysis. Models could be estimated successfully for the two longer distance bands; the recommended values for the shortest distance band were deduced from this.

Both own and cross elasticities were included in the estimating equations. In order to assist the estimation, constraints were imposed on the cross-elasticities using:

- diversion factor relationships: $e_{12} = e_{22} * \frac{Q_1}{Q_2} * d_{21}$, with d_{21} the proportion of passengers leaving ticket 2 when the price of ticket 2 rises that transfers to ticket 1. ITS have provided us with the diversion factors that have been used, shown in Table 2.2.

Table 2.2
Diversion Factors Used by ITS

To From	First	Full	Reduced	Apex	Stop travelling
First	-	0.60	0.20	0.00	0.20
Full	0.05	-	0.70	0.05	0.20
Reduced	0.00	0.50	-	0.10	0.40
Apex	0.00	0.10	0.40		0.50

- the Slutsky symmetry condition: $e_{12} = e_{21} * \frac{R_2}{R_1}$, with R the revenue associated with each ticket type.
- an assumption that the cross elasticity between first and reduced tickets is zero.

ITS have advised us that not all constraints have been used all the time. In some cases, a constraint was dropped if the model did not perform satisfactorily with the constraint.

The data shown in Table 2.1 indicate that the total price elasticity of demand for shorter distance journeys (20 to 100 miles) is slightly lower (less price elastic) than the price elasticity of demand for longer distance travel (100 miles and above).

In addition, the PDFH gives elasticity modifiers with which the ***total*** elasticities can be modified according to the price charged per mile, with the elasticities increasing (in absolute terms) for higher fares. These were based on the demand for all non-Apex tickets and have been estimated, pooling data across the three distance bands, according to the following formula:

$$f = a(£PM^b)$$

where:

f = elasticity for a given fare per mile

a = overall elasticity

PM = fare per mile

b = parameter driving the degree of elasticity variation according to the fare per mile. This parameter was constrained to be the same across all three distance bands

As shown in Table 2.3, the modified elasticities vary quite widely according to the average level of fares per mile.

Table 2.3
Variation of Fares Elasticities with Fare Level and Distance

20-100 miles		100-200 miles		Over 200 miles	
£ per mile	Elasticity	£ per mile	Elasticity	£ per mile	Elasticity
0.17	-1.07	0.16	-1.20	0.11	-1.30
0.15	-0.99	0.14	-1.10	0.09	-1.16
0.13	-0.90	0.12	-1.00	0.07	-1.00
0.11	-0.82	0.10	-0.90	0.06	-0.91
0.10	-0.78	0.08	-0.80	0.05	-0.81

2.1.3. London TCA and South East

The recommended elasticity values for London TCA and South East are contained in Table 2.4.

Table 2.4
Recommended Elasticity Values for London TCA and South East

To London		From London		Non London	
Seasons	Other	Seasons	Other	Seasons	Other
-0.3	-0.66	-0.6	-0.72	-0.6	-0.72

These values are taken from RIFF. In RIFF, elasticities are provided by journey purpose, and these have been converted into season and daily ticket elasticities by applying a commuting/business/leisure split for each ticket type. RIFF does not distinguish between “From London” and “Non London”, so that elasticities between these two groups are identical.

2.1.4. Non-London Long Distance (Without Full Set of Tickets)

This segment covers flows where first and full fare tickets account for less than 10% of the volume of ticket sales. For this segment, an overall elasticity of –0.9 is recommended. However, this elasticity depends on the fare per mile and ranges from –0.66 for low fares (£0.05 per mile) to –1.11 for high fares (£0.09 per mile).

The recommended elasticities are taken directly from the Handbook update demand analysis study.

2.1.5. Non-London Long-Distance (With Full Set of Tickets)

For this segment, an overall elasticity of –0.85 is recommended without differentiation according to the fare per mile paid. Flows in this segment largely occur on London routes but without a London origin or destination.

The recommended elasticity is taken directly from the Handbook update demand analysis study.

2.1.6. Non London Short Distance Flows (20 miles and less)

The recommended values for this segment are contained in Table 2.5.

Table 2.5
Recommended Elasticity Values for Non London Short Distance Flows (20 miles and less)

	PTE	Cross Boundary	Non PTE
Season tickets	-0.6	-0.6	-0.6
Prepaid off-peak tickets	-0.9	-0.9	-0.9
Daily tickets	-0.35	-1.0	-0.75

For daily tickets only, separate elasticities are provided according to the fare per mile charged. These values and the overall values for daily tickets have been taken from the Handbook update demand analysis study.

For season tickets, the recommendation for commuters contained in RIFF has been taken, whereas the RIFF-recommendation for leisure travel has been adopted for prepaid off-peak tickets.

2.2. Studies into Long-Term Elasticities

2.2.1. Railways

There have been two key previous studies that produced estimates of long-term fare elasticities for rail travel. These are:

- the study by Owen and Phillips (1987) on the characteristics of inter-city railway passenger demand; and
- the NERA (1999) study for OPRAF on forecasting passenger rail demand.

Owen and Phillips (1987) analysed the demand for rail travel on 20 London-based inter-city rail flows, using four-weekly data from 1973 to 1984. A partial adjustment model was estimated in which the number of single journeys between two stations was a function of GDP; average revenue per journey; a time trend, a dummy variable representing the introduction of High Speed Trains; a dummy variable representing coach deregulation; a dummy variable representing the introduction of air shuttles on some routes; seasonal dummy variables; and a lagged dependent variable. The model was estimated both for all journeys and for first and second class journeys separately. In the latter case, an additional fares variable was added so that both own and cross price elasticities could be estimated.

The fares elasticities estimated by Owen and Phillips are summarised in Table 2.6.

Table 2.6
Summary of Fares Elasticities Estimated by Owen and Phillips (1987)

	Short-term fares elasticity	Long-term fares elasticity
All journeys	-0.69	-1.08
1st class journeys	-0.67	-1.00
2nd class journeys	-0.81	-1.17

Source: Owen and Phillips (1987)

Owen and Phillips also estimated both GDP elasticities and time trends. In the case of GDP elasticities, nine out of their 20 estimates were not significantly different from zero. Similar problems affected their estimated time trends.

NERA (1999) undertook a study for OPRAF on forecasting passenger rail demand. The study consisted of a review of existing forecasting models; an econometric analysis of passenger rail demand; and the construction of a forecasting model to provide demand forecasts for each Train Operating Company and route over a 15 year period.

The econometric analysis examined the relationship between the total demand elasticity with respect to fare (estimated as total revenue across all ticket types divided by total journeys) and other explanatory variables, including GDP. Two estimation methods were used to derive elasticities:-

- static panel data estimation; and
- dynamic flow-by-flow estimation, incorporating a lagged dependent variable, similar to the approach used by Owen and Phillips.

Approaches to combining the panel and dynamic approaches were explored using the "Pool Mean Group" method.

A total of six panels were examined (the numbers of flows in each panel are shown in brackets):

- London-based Intercity flows (20);
- Short distance (London-based commuter services (10);
- Longer-distance London-based commuter services (16);
- Regional Inter-urban and non-London based Inter-city (25);
- PTEs (6 flows); and
- Other regional flows (10).

Data on journeys and prices (expressed as real average revenue per journey) were extracted from the CAPRI system for the purposes of the study and supplied by OPRAF. In a number of cases, it was necessary to correct the data for apparent serious under-reporting of journeys. GDP data (in real terms) were collected both at a national and regional level, but national figures were preferred as these are available more recently and at a higher frequency than regional GDP figures. Other variables such as generalised journey time, petrol prices, traffic speeds and revenue protection and privatisation effects were experimented with but found to be either insignificant or the source of estimation problems. These other variables have therefore generally not been included.

The results of estimating the dynamic models are shown in Table 2.7. The NERA-City estimates of the long term fares elasticities for London-based InterCity flows are considerably lower than those obtained by Owen and Phillips, whose results were shown in Table 2.4 above. NERA-City's estimates of the long run elasticity of rail demand with respect to GDP are also very substantially greater than those obtained by Owen and Phillips.

Table 2.7
NERA-City Long Term Elasticity Estimates
(mean group estimates for each panel)

Panel	Fare elasticity	GDP elasticity
London-based Intercity	-0.61	1.56
London commuter services	-1.12	1.67
Non-London based Intercity/Interurban	-0.97	1.20
Other regional	-1.06	1.62
PTE	-1.50	1.39

Source: NERA (1999)

When compared to the then PDFH (PDFH3), NERA's elasticity for London-based Intercity flows was rather lower than the average recommended by the PDFH. For London commuter services, by contrast, relatively high values were found. NERA provided a number of possible issues that might have affected the estimates, particularly in the case of

London commuter services, and indicated that these needed to be investigated. However, we did not rule out the case that our findings were accurate.

2.2.2. Local bus services

Dargay and Hanly (2002) examined the demand for local bus services in England, using annual data for 1986 to 1996 on bus patronage, fares and other relevant factors influencing bus use.

The analysis was undertaken at the county level, so that information on income and population for the individual counties could be used. Fares were approximated by average revenue per journey (including concessions). Bus vehicle kilometres per capita were used as the proxy for level of service.

The other explanatory variables used in the model were the per capita disposable income; population density; the percentage of pensioners in the population; and motoring costs. Data at the county level were used for all of these except motoring costs, for which only national data were available.

As the time period for the county data was too short to apply cointegration tests, a partial adjustment model was used. Two alternative specifications were examined, one with constant elasticities and one with elasticities related to the fare level.

The model was estimated by pooling the time-series data for the individual counties. Both a fixed effects and a random effects model were estimated, but the fixed effects model was found to be superior.

Two forms of the pooled model were estimated for each of the two specifications. In one form, all slope coefficients were constrained to be the same for all counties, implying identical elasticities between counties. In the other form, the coefficient of the fare variable was allowed to vary between regions. This resulted in a total of four model variants.

The results showed that the best-performing model variant was the one where the fare elasticity was allowed to differ between counties and was also dependent on the fare. Considerable variation in the fare elasticity across counties was found; ranging from 0 to –3.0 in the long run. For England as a whole, the Dargay-Hanly results indicate that the fare elasticity is likely to be about –0.4 in the short run and –0.9 in the long run. Demand is more price sensitive at higher fare levels, with elasticity value in the short/long-run ranging from 0.1/0.2 for the lowest fares to 0.8/1.4 for the highest fares.

2.2.3. Petrol

There exists a well-developed body of literature on the demand for automobile fuel. A meta-analysis by Espey (1998) used articles published between 1966 and 1997 which gave 363 estimates of short-run price elasticity and 277 estimates of long-run price elasticity.

An overview of the main trends in the literature is provided in Graham and Glaister (2002). Graham and Glaister note that although there is variation in the estimated elasticities of fuel demand, there are fairly narrow ranges within which the values typically fall. Short-term price elasticities are often between –0.2 and –0.3, whereas long-term elasticities tend to be between –0.6 and –0.8.

However, the authors also note that concern has been expressed in recent studies about the treatment of time-series in many earlier studies, particularly the lack of recognition of the non-stationary nature of the data. Most studies have simply estimated the demand for petrol as a function of the real price of petrol, real income per capita and a lagged dependent variable.

A number of studies have tried to take account of the non-stationary nature of the data by modelling the demand for petrol using the cointegration technique. These studies are summarised in Table 2.8.

Table 2.8
Summary of Petrol Price Elasticities Using Cointegration Techniques

	Short-term price elasticity	Long-term price elasticity	Adjustment coefficient
Bentzen (1994)	-0.32	-0.41	0.67
Eltony and Al-Mutairi (1995)	-0.37	-0.46	0.52
Samimi (1995)	-0.02 (insignificant)	-0.12	0.48
Ramanathan (1999)	-0.21	-0.32	0.28
Range typically reported in literature[1]	-0.2 to -0.3	-0.6 to -0.8	-

Table 2.8 shows that while most of the short-run elasticities found in the cointegration studies are in line with the values typically reported in the literature, the long-run elasticities are well below those reported in the major reviews of the literature. The authors of the studies where cointegration methods have been used typically state that this is at least in part because of the explicit treatment of the non-stationary nature of the variables. Graham and Glaister (***op. cit***) believe that the generality of the results is still open to question. They also suggest that many of the cointegration studies may have failed to take into account the possibility that price elasticities have grown over time as a result of increased fuel efficiency.

The table also shows the adjustment coefficients that were found in the studies that used cointegration methods. Adjustments coefficients indicate the proportion of the long-term demand adjustment change that occurs in the first year if demand drifts away from its long-

1 Source: Graham and Glaister (2002)

term equilibrium level. The studies show considerable variation in this proportion, with values ranging from 28 to 67 per cent.

3. REVIEW OF PREVIOUS TRENDS

3.1. Introduction

In this chapter, we review trends in demand and fare levels over the period from 1989/90[2] – to date.

We have organised our discussion around the four panels that we will use in our empirical work (see Section 4.4.1). These panels are:

- London Travelcard Area to and from Rest of Country;
- Non-London Long Distance;
- London Travelcard Area and South East; and
- Non-London Short Distance.

The first two panels are analysed in Section 3.2. Section 3.3 discusses the London and South East panel, whereas the Non London Short Distance panel (regional services) is covered in Section 3.4. In Section 3.5, we briefly discuss the impact of the Hatfield accident on rail demand.

3.2. Long Distance

In the Long Distance segment, we have two panels, one focusing on London flows and the other on non-London flows.

In addition to analysis of trends in total traffic and average fares in each of the long distance panels, we have also examined trends in demand and average revenue for the following broad ticket type categories:

- First: journeys in First Class on full fare tickets;
- Full: journeys in Second Class on full fare tickets;
- Reduced: journeys in Second Class on Saver, Supersaver, Apex and Day Return tickets; and
- Other: other tickets (both First and Standard Class), often operator-specific.

---

[2] The analysis of short-distance flows with significant season ticket volumes is based on so-called spread data, which means that the analysis must be undertaken over a shorter period, from 1990/91 – to date. See Section 3.3 below.

We discuss recent trends in each of the two panels below, starting with patronage and followed by average revenue.

3.2.1. Patronage

Figure 3.1 shows the number of journeys on our flows in the London TCA to and from Rest of Country panel. These are all medium to long distance inter-city flows.

Figure 3.1
Trends in Demand: London TCA to and from Rest of Country, 1989/90-2002/03[3]

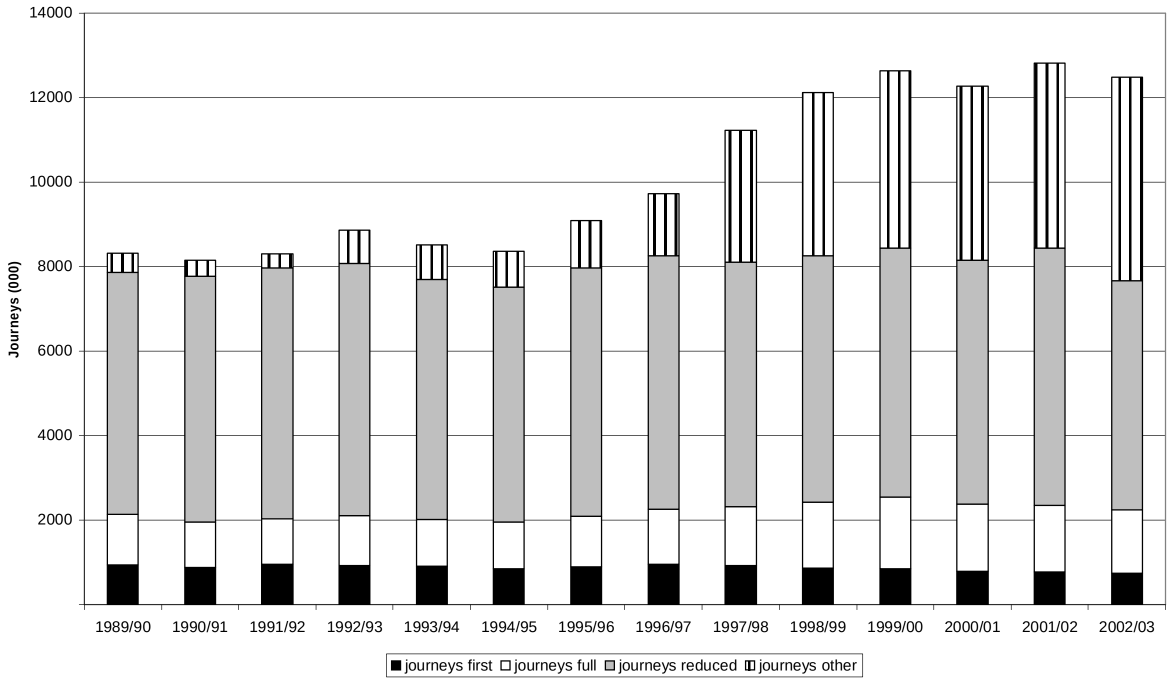


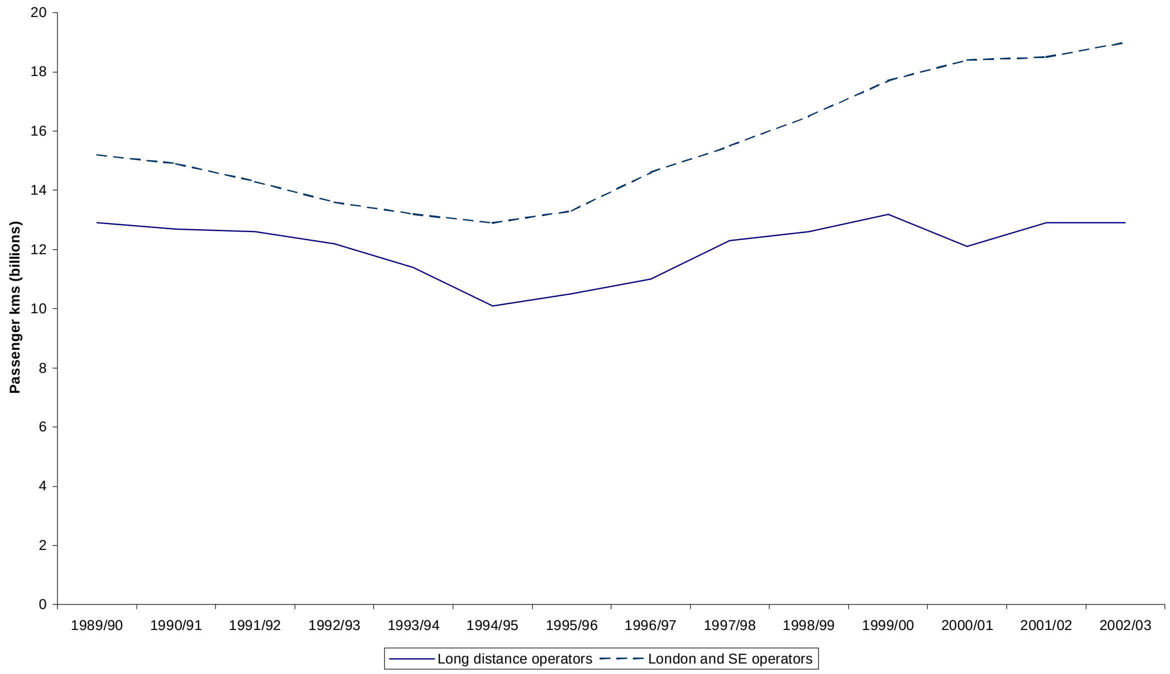
Figure 3.1 shows that the total number of journeys on longer-distance inter-city services to and from London has risen substantially, with all of the growth occurring in the period 1994/95 to 1999/00. The entire growth occurred in the "other" category which includes sometimes innovative, operator-specific tickets. The other ticket type categories remained more or less constant or declined somewhat. Since the average revenue from sales of "other" tickets was generally less than the average revenue from standard types of reduced fare ticket, the majority of "other" tickets would appear to represent sales of reduced fare products.

It is worth noting that the trend in long-distance journeys to and from London shown in Figure 3.1 is somewhat different from the trend in the total volume of passenger kilometres

[3] CAPRI data in the "other" category during 2000/01 are unreliable due to recording errors. For the purpose of this chart, we have assumed that travel in the "other" category fell between 1999/00 and 2000/01, as a result of the Hatfield accident and the subsequent disruption, by the same percentage as traffic in the "reduced" category.

carried on BR InterCity services and the successor long-distance TOCs, and shown in Figure 3.2.[4] Thus, between 1989/90 and 1994/95, the Transport Statistics series recorded a substantial fall in passenger kilometres, in line with the fall in passenger kilometres on London and SE operators (also shown in the figure). But this reduction is not reflected in the data for our panel shown in Figure 3.1.

Figure 3.2
Passenger Kilometres on Long Distance and London and SE Operators, 1989/90-2002/03



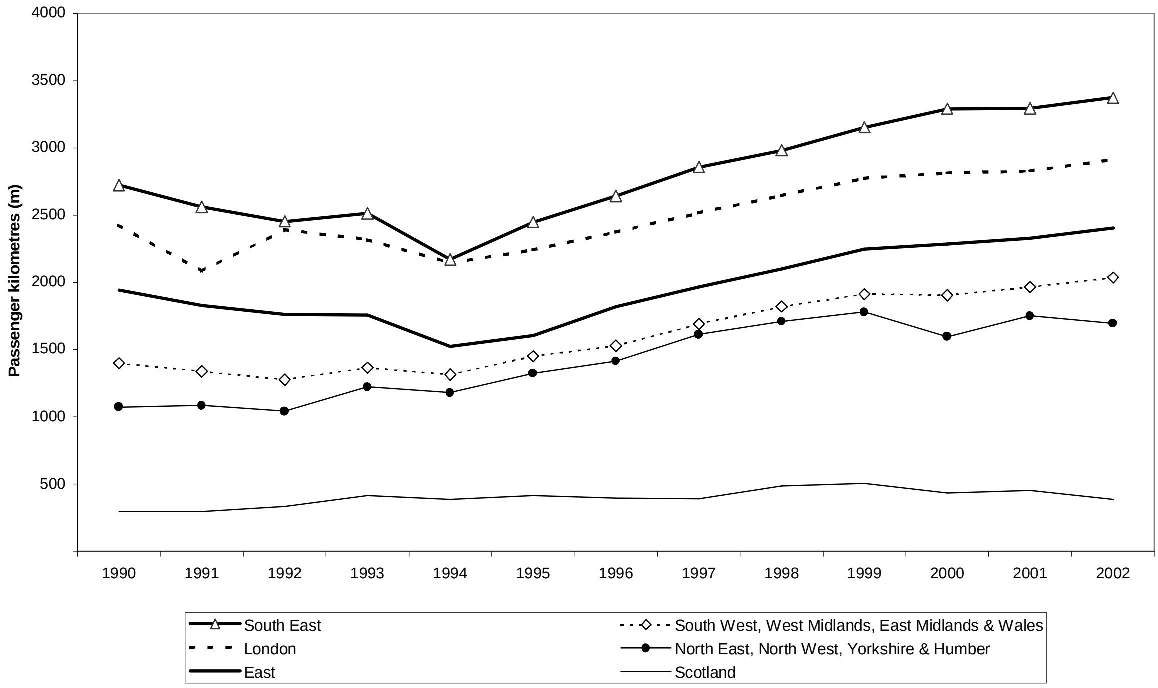
After investigation and with the assistance of SRA, we have established that the difference between the two series is due to the distinction between long distance travel on the one hand and the volume of traffic carried by long distance operators on the other. Our panel contains only medium to long distance flows and can therefore be regarded as being representative of trends in long distance travel. Long distance operators however carry a mix of long distance and shorter distance travellers, the latter mainly consisting of commuter journeys made on season tickets from stations such as Peterborough and Reading.

In Figure 3.3, passenger kilometres on services provided by long distance operators to and from London are broken down by region of origin or destination. It can be seen that most of the passenger kilometres on long distance operators are actually accounted for by journeys to and from London, the South East and East, all relatively short distance. Moreover, it is these short-distance categories that saw falls in patronage during the early 1990s reflecting the general reduction of commuting traffic into London. By contrast, traffic from the more

4 The data in Figure 3.2 are taken from Transport Statistics Great Britain.

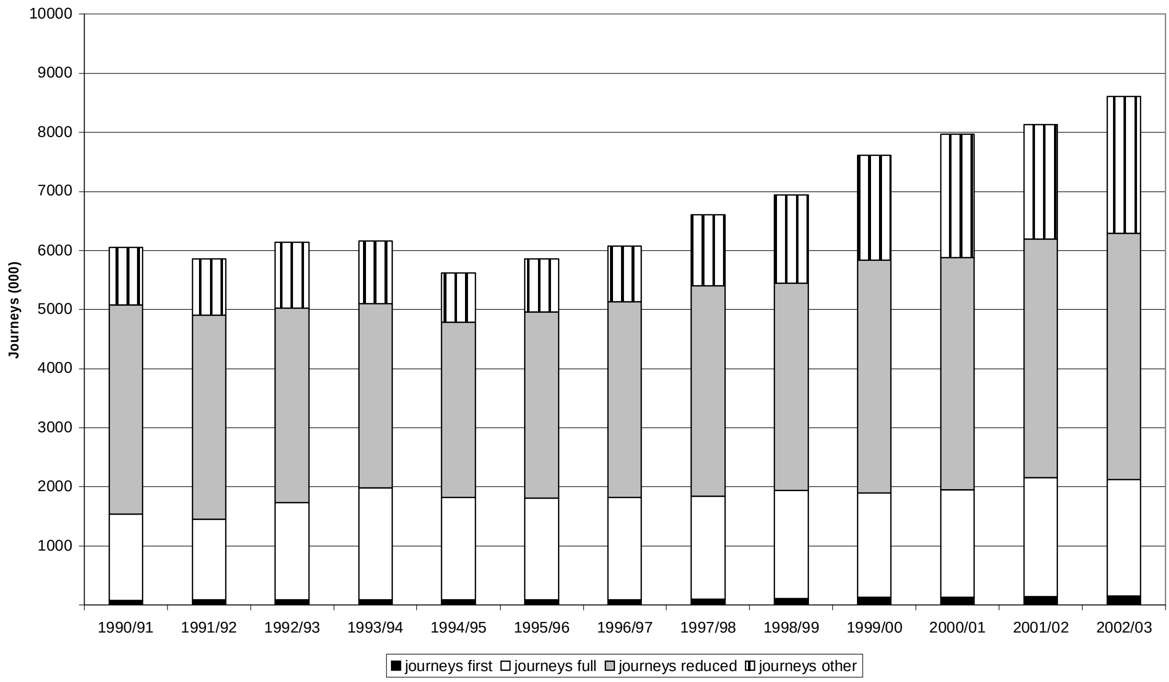
distant regions - the South West, the Midlands, Wales, the North and Scotland – remained broadly stable in the early 1990s or even grew somewhat.

Figure 3.3
Passenger Kilometres to and from London by Region, 1990-2002



The development of passenger numbers in the non-London long distance segment is shown in Figure 3.4. As in the case of the London long distance panel, this segment has shown strong growth after privatisation. It is again the "other" ticket type category that has grown most rapidly, although in this segment, travel in the "reduced" category has also substantially increased. Full fare tickets in second class have also increased in the last two years. First class travel on full fares in this segment accounts for a very small proportion of total ticket sales, though it did over the 12-year period increase by more than 100 per cent.

Figure 3.4
Trends in Demand: Non-London Long Distance, 1990/91-2002/03

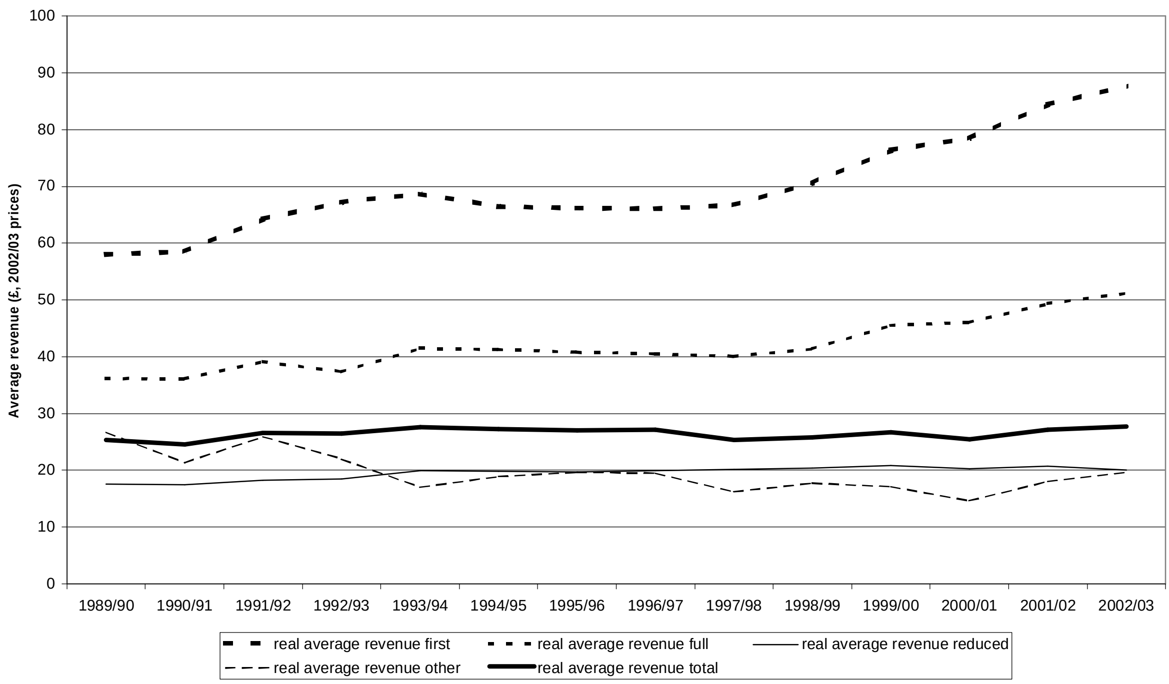


3.2.2. Average revenue

As shown in Figure 3.5, the average revenue per journey in the London TCA and Rest of Country panel has over the period from 1989/90 to date been relatively stable, on balance rising by about 10 per cent over the entire period. The limited changes in average real fares since privatisation disguise however wide differences in the behaviour of fully flexible tariffs and restricted availability discounted tariffs, the latter including both regulated Saver fares, and (unregulated) "other" fares. Full fare first and standard class tickets have risen sharply in real terms, as can be seen in Figure 3.5. By contrast, average revenue in the "reduced" category, has risen more slowly. Average revenue in the "other" category fell initially but has subsequently fluctuated around a more or less stable average.

Since privatisation, average revenues in the two premium ticket type categories have risen in real terms, while average revenues in the two reduced fare categories having been more or less stable. The reason why real average revenues in the entire panel have since privatisation also been stable is that the relative importance of the "other" segment, with low average revenues, has increased considerably (see Figure 3.1).

Figure 3.5
Trends in Real Average Revenue: London TCA to and from Rest of Country, 1989/90-2002/03[5]

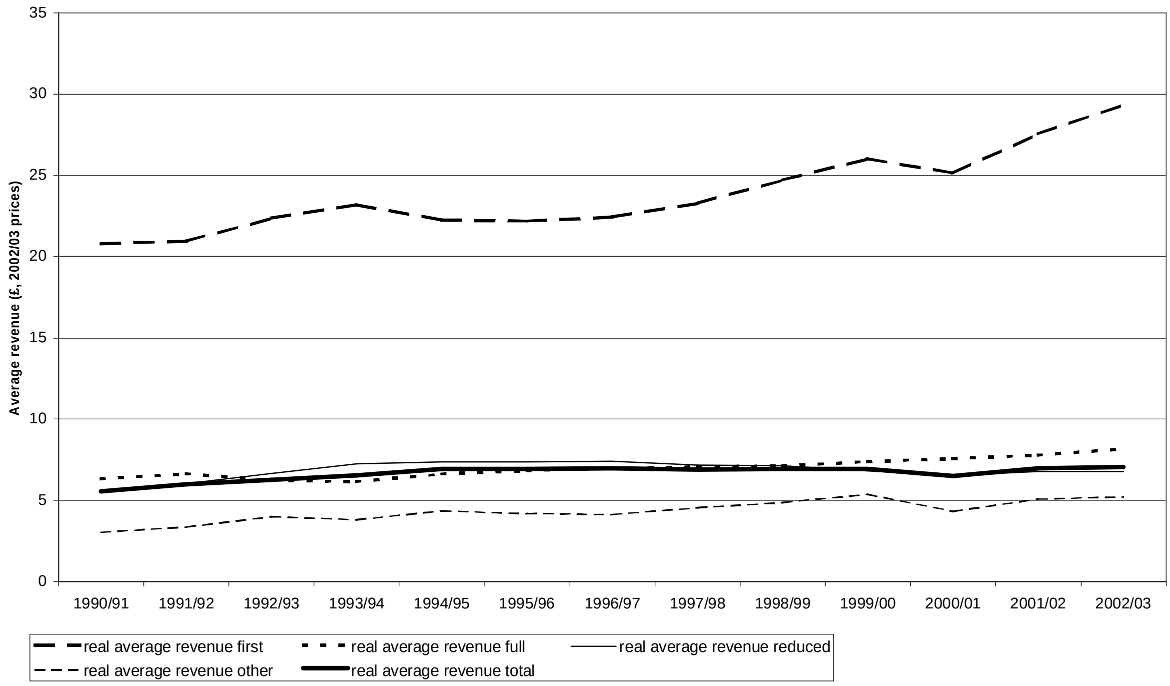


5 CAPRI data in the "other" category during 2000/01 are unreliable due to recording errors. For the purpose of this chart, we have assumed that travel in the "other" category fell between 1999/00 and 2000/01, as a result of the Hatfield accident and the subsequent disruption, by the same percentage as traffic in the "reduced" category.

In the Non London Long Distance panel, shown in Figure 3.6, the picture is somewhat different. Unlike in the London panel, real average fare levels in this panel have increased somewhat. As with the London-based flows, fully flexible first class fares have increased strongly, but the average real revenues for fully flexible, discounted standard class, and “other” fares, which together account for the large majority of all journeys in this market segment, increased slightly in the early 1990s, but have since remained broadly stable.

The increase in real average revenue across the panel also occurred mainly in the early 1990s, with a more constant level since then.

Figure 3.6
Trends in Real Average Revenue: Non London Long Distance, 1990/91-2002/03[6]



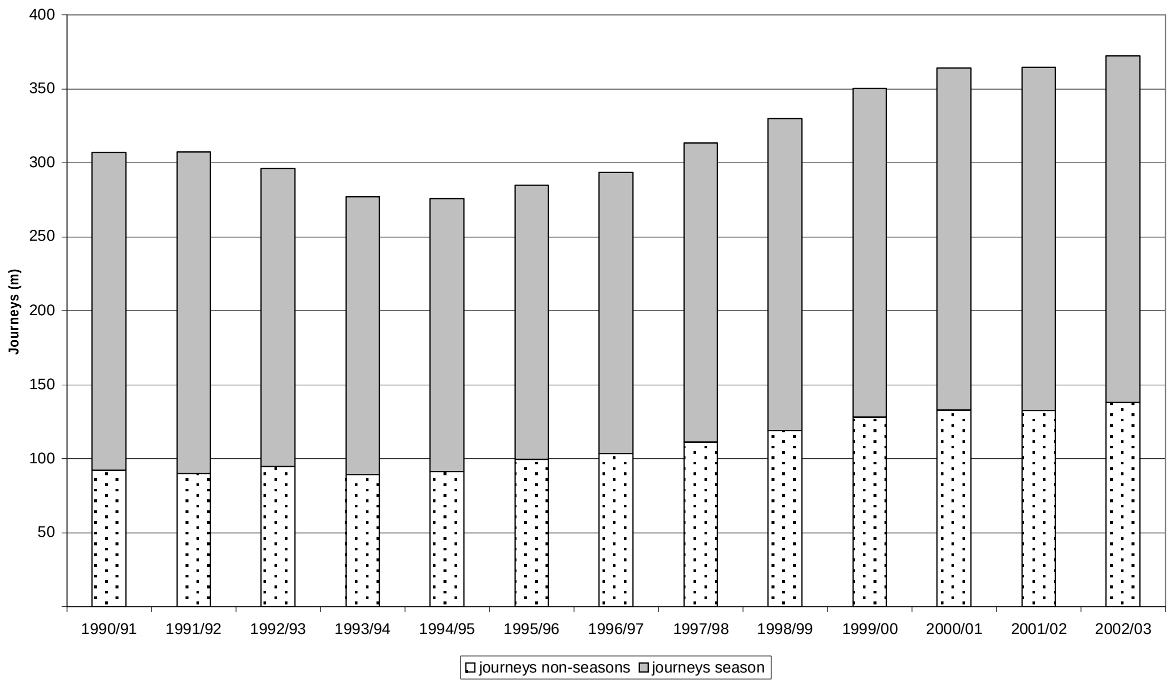
6 As can be seen, average revenue in the “reduced” category is higher than average revenue in the “full” category between 1992/93 and 1997/98. We believe this is due to the changes in the conditions attached to day returns, and the way these are recorded in CAPRI, on a number of flows in this panel over this period.

3.3. London and South East

3.3.1. Patronage

As we have already seen in Figure 3.2, the total volume of traffic on London commuter services fell sharply in the early 1990s, but then grew rapidly. Figure 3.7 below, which shows passenger numbers on flows within the London and the South East panel broken down into season and non-season tickets, replicates this pattern very closely.[7] We see that travel on season tickets fell particularly sharply in the early 1990s, but then recovered to a level of about 8 per cent above the previous (1990/91) peak by 2002/03. By contrast, travel on non-season tickets was stable during the early 1990s but then rose substantially. In 2002/03, the number of journeys on non-season tickets was 55 per cent above that in 1993/94. The combination of the rapid growth in non-season ticket travel, and slower growth in season ticket volumes has resulted in a significant fall in the share of travel on season tickets in the South East short distance panel flows, from around two-thirds in 1990/91, to around 55 per cent in 2002/03.

Figure 3.7
Trends in Demand: London and South East, 1990/91-2002/03



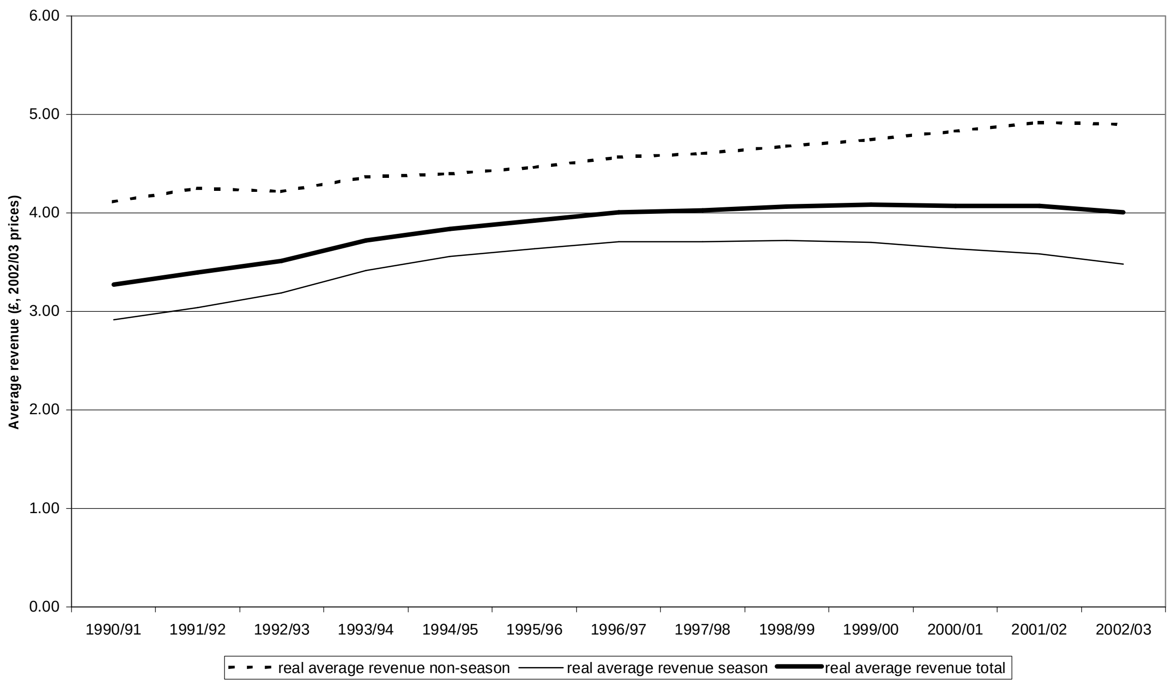
[7] In compiling this data series, we have "spread" travel on annual season tickets over the year following the period in which the ticket was purchased. This adjustment means that the first year of Capri data is lost, so that the analysis shows trends in volumes from 1990/91 onwards. As explained in Section 4.3.1.1 below, we have also used spread data in the econometric analysis of demand in the two short distance panels.

3.3.2. Average revenue

The development of average revenue in the London and South East panel is shown in Figure 3.8. Average revenue per journey in this panel rose sharply until the mid-1990s and more slowly afterwards. In recent year, it has started to fall somewhat.

Revenue trends in this panel are heavily influenced by the impact of economic regulation and the performance regimes. This applies in particular to season tickets, where real revenue rose up to 1998/99 but then fell markedly, so that it is now back at 1994 levels. By contrast, real revenue from non-season tickets showed a continued increase throughout almost the entire period. By 2002/03, it had reached a level about 20 per cent above that in 1990/91.

Figure 3.8
Trends in Real Average Revenue: London and South East, 1990/91-2002/03



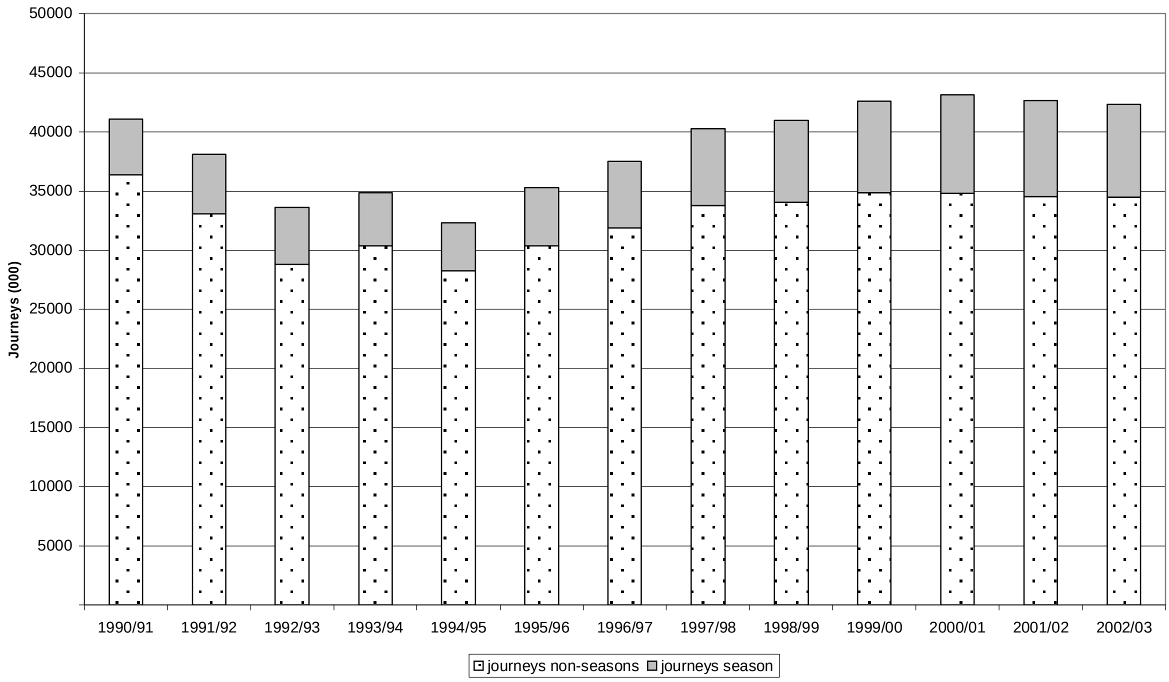
3.4. Regional

3.4.1. Patronage

In the Regional (NonLondon Short Distance) segment, the reduction in travel during the economic recession of the early 1990s, which also coincided with sharply increasing real fares (see 3.3.2 below) was even more marked than on services in the South East. As shown in Figure 3.9, both season and non-season ticket sales fell by about 20 per cent between 1990/91 and 1994/95. Season tickets quickly recovered and went on to grow spectacularly: in the peak year 2000/01, season ticket sales were more than twice as high as in 1994/95.

Since then, sales have again fallen somewhat. By contrast, travel on non-season tickets recovered much more slowly and did not reach the level of the previous peak in 1990/91.

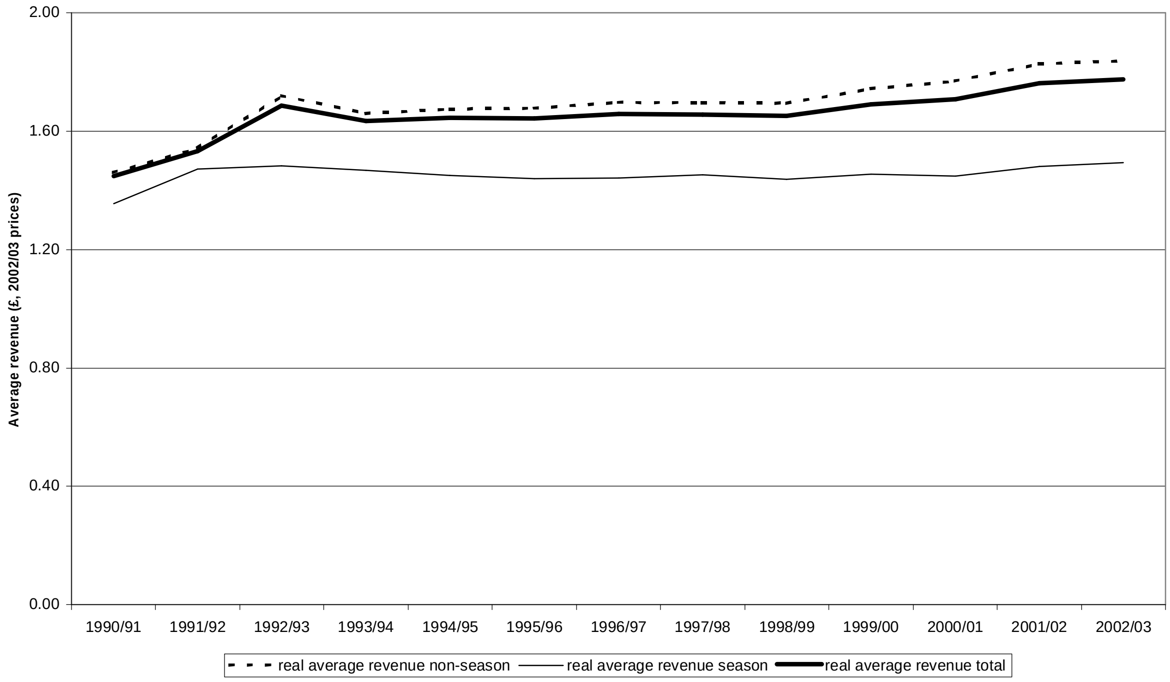
Figure 3.9
Trends in Demand: Non-London Short Distance, 1990/91-2002/03



3.4.2. Average revenue

Average revenue in the Non London Short Distance panel is shown in Figure 3.10. It rose sharply in the early 1990s, then remained stable for a number of years and rose further in the last couple of years. This trend is mainly caused by non-season tickets, which account for most of the revenue in this panel. Real average revenue from this category is now over 25 per cent above the level in 1990/91. Real average revenue from season tickets increased by about 10 per cent in the early 1990s, but has been stable over most of the remaining period.

Figure 3.10
Trends in Real Average Revenue: Non London Short Distance, 1990/91-2002/03



3.5. The Impact of the Hatfield Accident

The impact of the Hatfield accident, which occurred in October 2000, and the subsequent speed restrictions, is evident in several of the traffic volume and average revenue series discussed above. Figure 3.1 and Figure 3.2 each show declines between 1999/2000 and 2000/01, and Figure 3.3 shows reductions in passenger numbers on long distance flows, particularly between London and the North of England. These declines are striking because they followed a period of very strong traffic growth. Figure 3.5 and Figure 3.6 indicate temporary reductions in average revenue in a number of ticket categories. In particular, average revenue in the "other" category fell temporarily as a result of special offers by TOCs at the time service quality was recovering from the very severe disruptions that occurred in the immediate post-Hatfield period.

By contrast, the post-Hatfield disruptions appear to have had relatively little impact on shorter-distance flows into London and elsewhere, with some increases in traffic recorded between 1999/2000 and 2000/01 (Figure 3.7 and Figure 3.9).

In Section 5.5, we will report the result of our econometric analyses into the impacts of the Hatfield accident and the subsequent disruption.

3.6. Summary

The most important trends that we have noted in this Section and that will impact on our analysis are the following:

- In the two long distance segments (London and non-London), a substantial growth in the number of journeys has occurred, with all of the growth occurring in the "other" ticket category. This category is largely formed by operator-specific tickets, sometimes of an innovative nature and generally with low average fares.
- In the London TCA to and from Rest of Country segment, we have established an important distinction between long-distance travel and the volume of traffic carried by long distance operators, since the latter also includes substantial volumes of London commuter traffic. The difference was visible in the early 1990, when long distance travel remained stable but volumes carried by long distance operators fell.
- Although real average revenue in the two long distance segments has remained stable or increased slightly, fully flexible fares have generally increased sharply, whereas reduced fares have either fallen somewhat or become much more important in relative volume terms.
- In the two short distance panels, patronage fell in the early 1990s but has increased since then. Real average fares in these panels have increased, with less variation in the movements of individual fares than in the two long-distance panels.

4. EMPIRICAL METHODOLOGY

4.1. Introduction

The broad empirical approach we follow to estimate long-run fare elasticities is to construct and estimate a multivariate model of rail demand using econometric techniques suitable for the data available. In this chapter, we begin by setting out a generic framework for our empirical model of rail demand, which serves as a guide for the rest of this chapter. We then describe the data collected and assembled, discuss the issues surrounding the estimation of fare elasticities and present our econometric methodology.

The structure of the discussion is as follows:

- Section 4.2 presents a model of rail demand in its most generic form;
- Section 4.3 summarises the data that we have collected;
- Section 4.4 discusses the organisation of our analysis;
- Section 4.5 discusses estimation issues; and
- Section 4.6 sets out our econometric framework

4.2. A Model of Rail Demand

Our model of rail demand follows directly from previous studies, as reviewed in Chapter 2 (eg NERA-City (1999), Jones and Nichols (1983) etc). The factors that are considered to be important determinants of rail demand are: real income, economic activity, fares, generalised journey time, service quality, price and quality characteristics of alternative transport modes, and the time of year. Putting these factors together, the model of rail demand can be expressed in its most generic form as:

$$D_i = f(Y, EA, P_i, P_a, GJT, SQ, AM, T) \qquad \text{(i)}$$

Where:

D_i	is the level of demand for ticket type ***i***;
Y	is real income;
EA	is an index of economic activity;
P_i	is the price of ticket type ***i***;
P_a	is the fare for other ticket types for the same rail journey
GJT	is generalised journey time;
SQ	is the quality of service the passenger can expect;
AM	is a set of performance (price, journey time, etc) factors for alternative transport modes; and
T	is a seasonal effect.

The precise formulation of our empirical model has been shaped by the dataset we have been able to construct and is discussed in Section 4.6. First, we describe the data we have collected on each of the factors in the model.

4.3. Data

A number of sources were drawn from to assemble the dataset used for our analysis. Table 4.1 provides a summary of the main data sources for each variable. In the subsections that follow, we discuss the data for each variable in detail.

Table 4.1
Summary of Data Sources

Variables	Data source
Demand (D_i)	CAPRI, Number of journeys
Price, for all ticket types for the same rail journey (P_i , P_a)	CAPRI, Revenue per journey
Real Income (Y)	Office of National Statistics, National GDP at 1995 prices
Generalised journey time (GJT)	SRA, Planned average timetable journey time plus half of the time between trains
Service quality (SQ)	SRA, Punctuality
Economic activity (EA)	ONS, (i) Synthetic index of economic activity based on GDP, and (ii) Unemployment rate
Attributes of alternative transport modes (AM)	(i) DTI, Petrol prices, and (ii) DfT, Vehicle kilometres

4.3.1. Demand and prices

Rail passenger demand data have been drawn from the Computer Analysis of Passenger Revenue Information (CAPRI) database. This database contains four-weekly data on the number of tickets sold and the revenue associated with each ticket type for each journey. The sample we have drawn contains 97 individual flows and includes all four-weekly periods from April 1989 to March 2003.

The CAPRI data are suitable for the present analysis due to both the length and periodicity of the available time series, and the level of disaggregation of the data. In particular:

- there are observations for 182 time periods for each flow between April 1989 to March 2003; and

- for each flow, the data include information on both the direction of the journey (i.e. reverse and outward) and for 14 ticket types.[8]

The price variable we use for the analysis is the average revenue, defined as total revenue divided by the number of journeys for each ticket type or grouping of ticket types. This variable is converted to constant 1995 prices using the RPI series drawn from Office of National Statistics.

4.3.1.1. Adjustments to CAPRI data

A number of adjustments were made to the CAPRI data prior to analysis:

- Some periods were omitted due to strike activity. During these periods it is possible that the number of ticket sales represented the level of capacity rather than the level of demand.[9]
- An adjustment factor has been applied to the data in some periods to take into account the fact that these periods contained more or less than 28 days. These periods are in either the first or the 13th period in a year. The adjustment factor we applied to both journeys and revenue was simply 28/N, where N is the number of days in the period
- Season ticket sales, which are recorded in the period in which they are sold, have been smoothed by spreading the associated journey numbers over the period of validity of the ticket. This was necessary because the traditional peak in the sale of annual season tickets in December/January, prior to the annual fares increase, has reduced considerably in recent years. Fares regulation has limited the annual increases in the prices of season tickets or even produced nominal decreases. As a result, there is now less of an incentive to renew annual season tickets immediately before the date of a fares increase. The fact that the seasonality pattern is not constant over the years meant that it was not possible simply to take account of the seasonal variation by using seasonal dummy variables.

4.3.1.2. Data cleaning

In several instances, it was clear that the raw CAPRI data severely misreported the number of journeys and / or revenue in a period. These cases were identified by inspecting the time series of individual flows and were removed from the dataset. Some of the most severe cases of misreported demand data were observed in the periods immediately following the Hatfield rail crash.

---

8 See Table 4.3 on page 36 for a complete list of the different ticket categories that are included in the CAPRI database.

9 This applies to the first period in 1991/92, and the third, fourth, fifth and sixth periods in 1995/96.

4.3.1.3. Composition of CAPRI flows

For the majority of flows in the sample supplied to NERA for the period 1990 to 1999, the CAPRI data measured ticket sales between groups of stations, rather than between individual stations. For example, the origin and destination points have names such as "Kent Link", "Croydon Lines", "Sutton Lines", etc. The groups contain anywhere between 1 and 10 individual stations, and these are in some cases spread over a large area. Although the CAPRI data supplied to NERA for the period 1999 to 2003 are measured for flows between individual stations, in order to obtain a consistent series for the whole period, the 1999-2003 flow level data were aggregated into the same groups.

The fact that the data do not represent individual flows could, in theory, introduce a bias on the estimates of demand elasticities. This would be the case if the composition of demand across flows within groups varies over time. The measure of price that we use in our analysis, average revenue, may rise or fall purely due to changes in the proportion of flows associated with individual stations within the group and hence changes in this variable may not reflect real price changes.

The effects of real price changes could thus be confounded by changes in average revenue, caused by compositional changes in demand. This would lead to a price elasticity estimate for the group that is biased downwards, ie lower than the weighted average of the true price elasticities of demand for the individual flows.

To assess the extent to which demand compositional changes within groups would be likely to affect estimated price elasticities of demand, we constructed and ran a simple simulation model using artificial observations on prices and journeys. We found that the effect on the price elasticity estimate in this model of fairly significant compositional changes was actually relatively minor.[10] We also examined changes in the distribution of ticket sales between stations in a sample of cases on London commuter flows, where we thought compositional effects might be most serious, because of the relatively high potential for switching between stations. We found that changes in the composition of the flows had, in fact been very small.

---

10 The likely size of compositional effects was evaluated by estimating a simulation model using data at the aggregate level when the underlying disaggregated data was subject to random variations in composition. These variations in composition were generated by fares that followed a random walk process, which then had an impact on journeys through the disaggregated demand elasticities in the model. The model was calibrated to give reasonably large but plausible variations in the relative number of journeys.

A total of 200 observations were generated for each realisation of the random processes.

Regression estimates were then derived on the basis of the aggregate journeys and aggregate fares for several realisations of the data. The resulting estimates were invariably close to the weighted average of the true elasticities for the different journeys.

On the basis of both the simulation analysis and the results of the empirical analysis, we concluded that compositional effects were unlikely to be a significant source of bias in our estimates of price elasticities.

4.3.2. Real income

The measure of real income in our dataset is UK GDP at 1995 prices, drawn from the Office of National Statistics. The UK GDP time series covers the entire sample period and is measured on a quarterly basis. The data have been converted into four-weekly periods using interpolation of the quarterly series.

We considered the possibility of examining regional GDP figures, but we found that these data were highly correlated with UK GDP for the time period in question. Regional GDP data are also available on an annual basis only. Given that our analysis is based on much more disaggregated time periods, we did not pursue this option further.

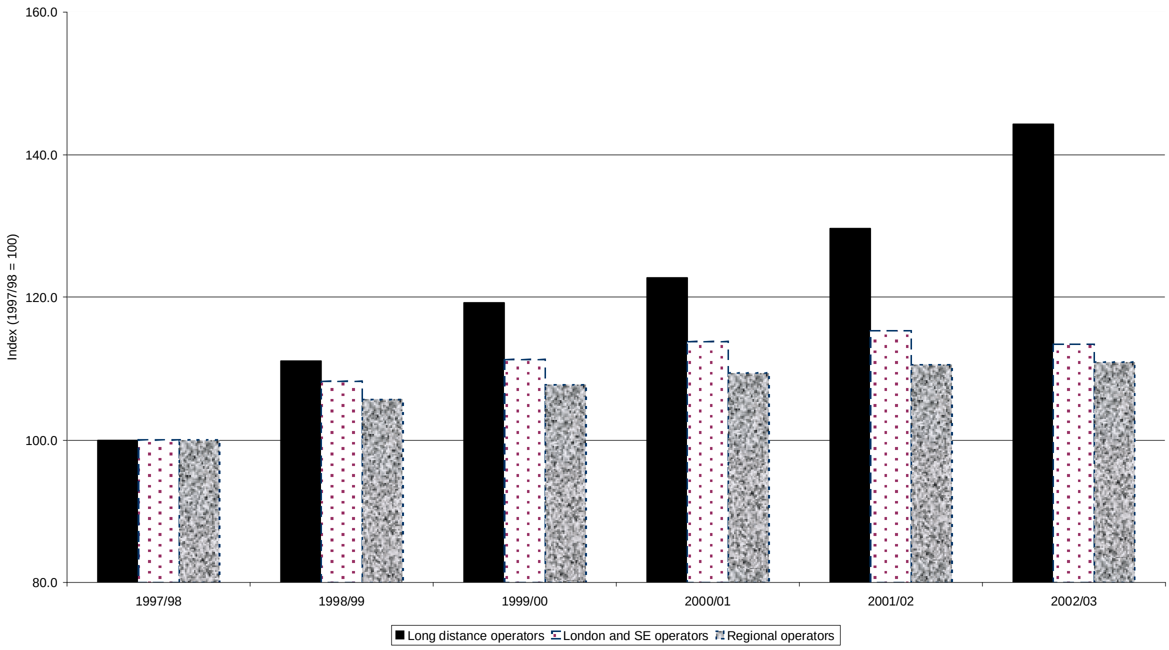
4.3.3. Generalised journey time

Generalised journey time (GJT) consists of the average timetabled journey time plus half of the time between trains.[11] The importance of GJT in influencing rail demand was shown after the Hatfield accident in October 2000. The speed restrictions that were imposed after the accident produced large increases in journey times, and, together with the overall reduction in service reliability, resulted in substantial reductions in demand on some parts of the network. Long distance markets were by far the most serious ones affected.

In addition, GJT is a function of the number of train kilometres scheduled to run on the network. More train kilometres implies higher frequencies, which in turn implies lower average waiting times. As can be seen in Figure 4.1, timetabled train kilometres have increased substantially in recent years, particularly in long distance services. In 2002/03 almost 45 per cent more train kilometres were scheduled for long distance services than five years earlier. In the two other segments, a more moderate growth occurred of between 10 and 15 per cent.

[11] Half of the time between trains is a commonly used estimation of the waiting time. This proxy is based on the assumption consisting of considering the travellers' arriving time at the stations as distributed according to a constant function of density.

Figure 4.1
Timetabled Train Kilometres by Sector, 1997/98-2002/03
(1997/98 = 100)



Source: National Rail Trends

We obtained data on GJT from SRA. Unfortunately the data do not capture the temporary increases in GJT post Hatfield, since these are based on ***planned*** annual timetables as opposed to ***actual*** timetables.

We attempted to obtain GJT data based on actual timetables in order to be able to model the effects of GJT on demand in the post-Hatfield period, however after discussion with SRA, it was decided that this was not appropriate since timetables were changing very frequently during this period often without advance communication to passengers. As a result, the GJT based on actual timetables could not have been considered by the passengers prior to choosing to purchase a ticket.

Because of these problems, we have modelled the impact of the Hatfield accident using dummy variables. Our approach is described in more detail in Section 4.5.1.

4.3.4. Service quality

Service quality in the context of rail journeys refers to the punctuality and reliability of train services. Figure 4.2 provides an overview of service quality in each of the three main market segments, showing in particular the impact of the Hatfield accident on punctuality levels. It can be seen that all segments were severely affected, most notably long distance services. Since the accident, performance levels have recovered somewhat in each of the segments but remain at low levels.

Figure 4.2
Percentage of Trains Arriving on Time, 1997/98-2002/03

Source: SRA National Rail Trends

For the purpose of this study, SRA provided detailed service quality data for each Passenger Charter Service Group. The two service quality measures are defined as follows:

- ***punctuality*** **-** the proportion of trains that arrive at their destination within five or ten minutes of their scheduled arrival time; and
- ***reliability*** **-** the proportion of planned train services that actually run.

The periodicity and length of the data on punctuality and reliability are appropriate for the purpose of the analysis as they have been recorded on a four-weekly basis from 1989/90 to 2003/04. However, a sizeable proportion of the data series were missing. Given the benefits of using as many observations as possible in the estimation process, we have interpolated the missing data using an average of the neighbouring non-missing records.

We have concerns about the usefulness of both measures of service quality as explanatory factors in our demand model. First, the reliability data contain very little variation within flows over time and, as a result, the data are unsuitable for econometric analysis. In addition, the punctuality data were based on actual timetables over the post-Hatfield period, and since timetables changed regularly, the punctuality data exhibit substantial variation that is unlikely to accurately measure the service quality as perceived by passengers. For example, when timetables switched back to normal, there were substantial drops in punctuality even though the service quality was in fact improving.

4.3.5. Economic activity

The level of cyclical economic activity is expected to affect the demand for rail travel independently of the level of GDP. An economic downturn, although consistent with stable GDP, could cause significant reductions in company profits, new business etc., and a reduction in consumer confidence, leading to a reduction in the demand for both business and non-business rail travel.

We considered two measures of the level of economic activity:

- an index based on the ratio of GDP to an estimated trend level of GDP
- the level of unemployment in both Great Britain and London.

We constructed a synthetic index to measure economic activity that consists of a ratio between the actual level of GDP, and an estimate of the GDP trend level.[12] The index is intended to reflect the ***deviation from growth trend*** or, in other words, to capture the cyclical component of the GDP.[13, 14]

The second possible way to measure the level of economic activity is the unemployment rate in both Great Britain and London. The data, which have been extracted from the NOMIS database of the Office of National Statistics (ONS), have been estimated with the claimant count records. The data are on a monthly basis but have been adjusted to a four-weekly basis to be consistent with the CAPRI demand data.

Of these, we have eventually used the economic activity variables which proved significant in some panels (see Section 4.6.3). Since the economic activity variable only captures the deviation from trend growth in GDP, we have in addition to the economic activity variable also included GDP as explained in Section 4.3.2 above.

4.3.6. Attributes of alternative transport modes

The price and service levels of non-rail modes of transport can be expected to affect the level of demand for rail travel where there are possibilities for substitution. The principal substitutes for rail travel are cars and buses/coaches. The collection of data on these transport modes proved a difficult task due to the fact that data on the attributes of alternatives modes are usually available, if at all, only nationally and/or annually. By

---

12 The estimation of the GDP trend was based on a method consisting of regressing GDP data against time during a period that starts and finishes at comparables moments of the economic cycle (i.e. peaks during the second and third quarters of 1973 and 2000 respectively).

13 It is worth mentioning that a similar synthetic index was used in previous studies and it proved to be adequate and significant in the data analysis (see for instance, ***Ian S. Jones and Alan J. Nichols, The demand for inter-city rail travel in the United Kingdom, JTEP, May 1983***)

14 The index is highly negatively correlated (-96 per cent) with the national unemployment rate.

contrast, the demand data in our sample are highly disaggregated both geographically and temporally.

We have selected a number of proxies that roughly reflect the generalised cost of the road transport at a national level. The selected proxies are:

- petrol prices; and
- vehicle kilometres on both motorways and non-built up major roads.

The effect of variation in petrol prices on demand has been examined using a monthly index of the national average price of premium-unleaded petrol, obtained from the Department of Trade and Industry. The time series covers the full sample period. We have converted the price data to 1995 prices, and have adjusted the monthly time series to convert them into four-weekly periods.

The effect of changes in road congestion on demand has been examined using an annual index of vehicle kilometres travelled on UK motorways and major roads as a proxy. The index was obtained from the Department for Transport. We have adjusted the annual data to convert them into four-weekly periods.

4.3.7. Seasonal variation

Seasonal changes in demand behaviour have been examined through the use of 12 dummy variables for the 13 four-weekly periods in a year.

4.4. Organisation of Analysis

4.4.1. Data panels

In line with the organisation of recommended elasticities in the PDFH, our analysis of rail passenger demand focuses on four distinct market segments. The individual flows in our sample are grouped into panels corresponding to these market segments. The numbers of flows in each panel are shown in Table 4.2. A full list of the flows in each panel is contained in Appendix A.

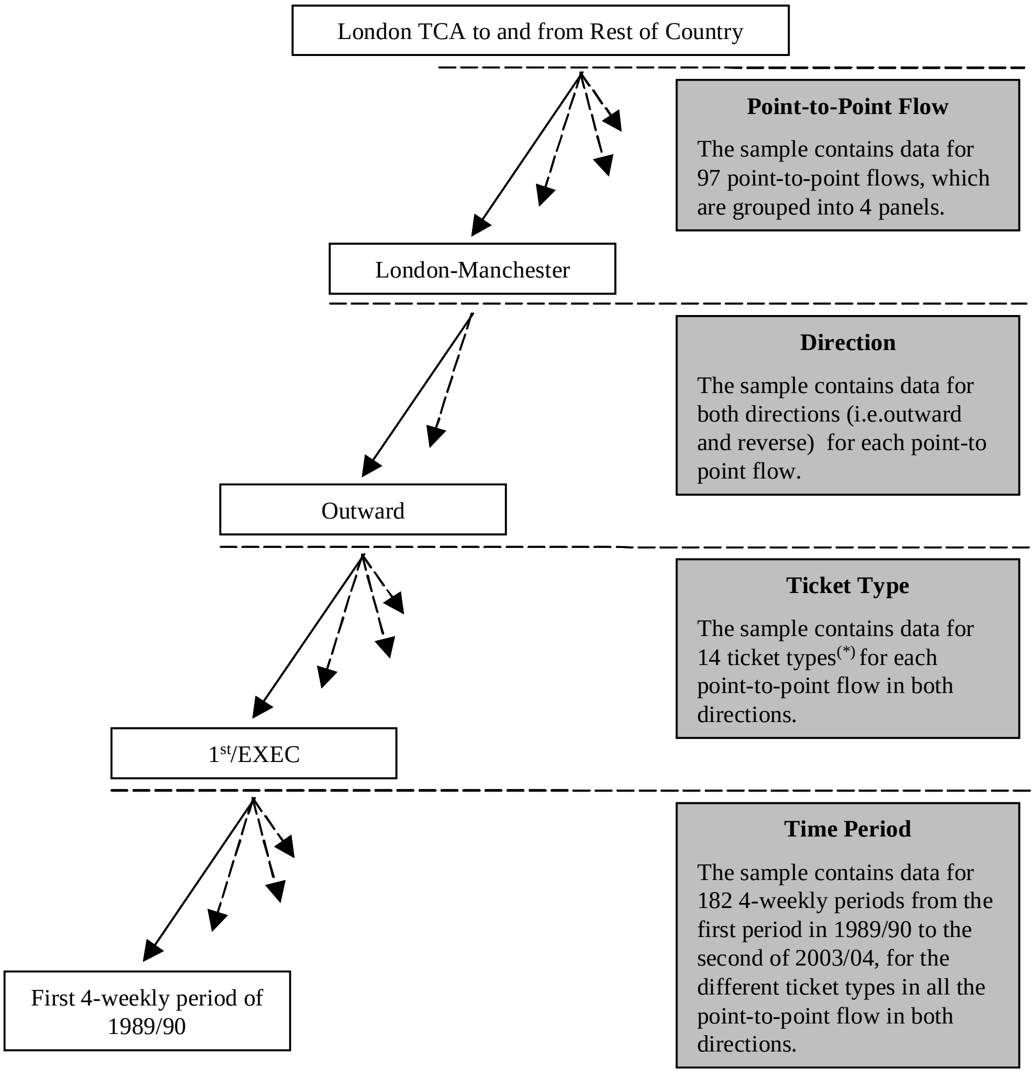
Table 4.2
Data Panels

Panel	Number of flows
London Travelcard Area to and from Rest of Country	20
Non-London Long Distance	35
London Travelcard Area and South East	26
Non-London Short Distance	16

4.4.2. Level of aggregation by ticket type

The CAPRI data we obtained for each flow contain ticket sales information for 14 product codes and for Outward and Reverse directions. An illustration of the potential level of disaggregation is provided in Figure 4.3.

Figure 4.3
An Illustration of the Level of Disaggregation in the Data Set



(*) The ticket types are (1) 1st/EXEC, (2) STD SINGLE, (3) STD/OPEN RTN, (4) SAVER, (5) SUPERSAVER, (6) NETWWK AWYBRK, (7) APEX, (8) DAY RETURN, (9) ANNUAL S/T, (10) WEEKLY S/T, (11) OTHER SEASON, (12) SLEEPER SUPP, (13) CHDY SINGLE, (14) OTHER

This in theory enables us to estimate elasticities for all the classes reported in the PDFH. In practice however, we consider that more reliable estimates may be obtained using data aggregated over all ticket types.

The main results we present in this report are derived from data on ticket sales aggregated over all product types. However, in line with the reported elasticities in the PDFH, we also estimate demand equations using disaggregated ticket sales. The approaches we follow and the issues involved in these two parts of our analysis are discussed below.

4.4.2.1. *Aggregate ticket sales analysis*

The aggregation of ticket sales data raises problems concerning the effect of changes in the composition of sales across the different ticket categories over time. Figure 3.1 in Chapter 3 (page 14) shows that the level of ticket sales for London based intercity flows in First class, Standard class and Reduced fare categories has remained relatively constant over the sample period. Ticket sales in the "Other" category meanwhile have increased substantially over this period, mainly due to the introduction of special operator specific tickets at lower fares than the Reduced category.

If there is growth in cheaper journeys, perhaps because of a decline in the effectiveness of coach competition, the aggregate data will show a simultaneous fall in average revenue and a rise in demand. There would be an apparent "price effect" even if there had been no actual change in the price of individual tickets.

It is possible to overstate this problem, however. To the extent that the increase in low fare tickets is a result of increased availability of these tickets then the observed fall in price is a genuine effect and will not lead to any obvious problems. Furthermore, for the purposes of financial modelling it is the revenue implications that matter.

4.4.2.2. *Disaggregate ticket sales analysis*

We have estimated demand equations disaggregated by ticket type categories for each of the four panels. Table 4.3 shows how the data in each panel have been disaggregated into groups of ticket types.

Table 4.3
Disaggregation by Ticket Type

<table>
<tr><th rowspan="2" colspan="2"></th><th colspan="4">Panels</th></tr>
<tr><th>London TCA to and from Rest of Country</th><th>London TCA and South East</th><th>Non-London Long Distance</th><th>Non-London Short Distance</th></tr>
<tr><td rowspan="14">Ticket types</td><td>1ST/EXEC</td><td rowspan="3">First+Standard</td><td rowspan="8">Other</td><td rowspan="3">First+Standard</td><td rowspan="8">Other</td></tr>
<tr><td>STD SINGLE</td></tr>
<tr><td>STD/OPEN RTN</td></tr>
<tr><td>SAVER</td><td rowspan="11">Reduced+Other</td><td rowspan="11">Reduced+Other</td></tr>
<tr><td>SUPERSAVER</td></tr>
<tr><td>NETWWK AWYBRK</td></tr>
<tr><td>APEX</td></tr>
<tr><td>DAY RETURN</td></tr>
<tr><td>ANNUAL S/T</td><td rowspan="3">Seasons</td><td rowspan="3">Seasons</td></tr>
<tr><td>WEEKLY S/T</td></tr>
<tr><td>OTHER SEASON</td></tr>
<tr><td>SLEEPER SUPP</td><td rowspan="3">Other</td><td rowspan="3">Other</td></tr>
<tr><td>CHDY SINGLE</td></tr>
<tr><td>OTHER</td></tr>
</table>

The groupings of ticket types have been chosen to be broadly in line with the level of disaggregation presented in the PDFH, reported in Section 2.1 of this report. However, because of changes in the composition of ticket sales in the period since privatisation, and restrictions in the accessibility of data in the CAPRI database, we have not attempted to derive estimates for precisely the same groupings of ticket types that are reported in the PDFH.

Long distance panels (London and non-London based)

The PDFH presents fare elasticities for First, Full and Reduced ticket types for the London TCA to and from Rest of Country market segment. It presents elasticities only for the aggregate set of ticket types for the Non-London Long Distance market segment.

As discussed in Section 3.2.1, one of the key areas of growth in the long distance market segments is in sales of tickets that fall into the "Other" category. This category contains operator specific tickets, which mainly have lower fares than even the Reduced ticket types.

Whilst it is understood that not all tickets in this category are reduced tickets, it was not possible to split the allocation of the "Other" tickets between the First, Standard and Reduced categories. Instead, since it is our understanding that the vast majority of these tickets are reduced fare tickets, and because the average revenue for these tickets is below that of even the Reduced category, we have combined Other tickets with the Reduced category.

In the inter-city markets, our approach has been to combine First and Standard full fare tickets into a single group for the panels concerned. This group is characterised by the twin features that choice of journey time is fully flexible and that no advance booking is required prior to travel. In contrast, tickets within the Reduced+Other ticket category will only be available outside specified peak periods and will often need to be booked in advance.

At a theoretical level, the choice between "First+Standard" and "Reduced+Other" product bundles can be characterised as a high level decision on when to travel. Choice within each of the two groups is then a lower level decision contingent on the prior decision between product groups. The lower level choice may be more efficiently investigated using Stated Preference approaches given the problems, such as highly correlated variables, that hamper investigation using time series econometric methods.

Short Distance Journeys

In the case of the "London TCA and South East" and "Non London Short Distance" panels, the ticket types have been aggregated into the two categories "Seasons" and "Other" as shown in Table 4.3. In the past, these market segments comprised two well-differentiated sub-markets, with season ticket travel being closely associated with journeys to work, and travel on other types of ticket being associated with more "optional" types of travel. Although we would expect that most travel on season tickets is work-related, changes in working patterns, particularly the greater incidence of part time and from home working may mean that an increasing proportion of work-related travel is no longer undertaken on season tickets. The demand characteristics of the two types of ticket may therefore be somewhat less differentiated than they were formerly. In view of this, we have in our analysis attempted to estimate cross-elasticities between the two categories of ticket.

The analysis of the London TCA and South East panel further distinguishes between inbound and outbound traffic, the former representing ticket sales made at the suburban station, by rail users travelling in to London, and the latter sales made at the Central London end for travel out from the centre. The volume of inbound sales is approximately 10 times as large as the volume of outbound sales. Inbound and outbound London commuter traffic will be analysed separately in particular to assess whether outbound journeys are less "captive" to rail than inbound journeys, for example, because of the greater availability or lower cost of parking at the trip end, and hence the closer substitutability of car for rail mode. Perhaps for this reason, the PDFH currently recommends a higher price elasticity of demand for outbound than for inbound travel.[15]

4.5. Estimation Issues

4.5.1. Modelling the effect of the Hatfield crash

One of the key areas of concern in selecting the set of explanatory variables was how to model the effect of the Hatfield crash on demand. Journey times increased significantly across a large part of the network in the post-Hatfield periods due to both decreases in frequency of service and speed restrictions. The crash and the accompanying media coverage also raised passenger concerns over the safety of the network.

In theory, this period presents a good opportunity to derive estimates of the effects of increases in journey times on demand because of the variations in journey times. However, the generalised journey time data we collected were based on planned timetables and not the actual timetables, which were substantially different for some flows during the post-Hatfield period. As a result, the data were unsuitable for modelling the effect on demand of the changes in journey times and service frequencies. The service quality data we collected were similarly unsuitable to capture the effect of the Hatfield crash since the data did not reflect the quality of service that would have been perceived by the passenger during this period (see Section 4.3.4).

The approach we have adopted to capture the effects of the Hatfield crash on demand has been to use dummy variables for the post-Hatfield periods. The use of dummy variables allows us to quantify the magnitude of the Hatfield effect as a proportion of normal demand

[15] In addition, inbound and outbound traffic are subject to different types of regulation. Inbound tickets are all regulated in baskets and were constrained with the RPI-1 policy. Outbound tickets have been subject to less regulation and interact less with the Travelcard Area.

levels while controlling for the explanatory variables in the dataset. We tested a number of sets of dummy variables to examine the Hatfield effect with alternative groupings of the four-weekly periods after the Hatfield crash.[16]

4.5.2. Simultaneity

A potential issue to take into account when estimating demand models is the nature of the supply side response to demand conditions. In the current context, the issue of concern is whether or not the price of the rail journey is set by the train operator independently of the level of demand, and if it is not, whether the price responds to demand shocks in the same four-week period. If the operator does set price in response to demand conditions within the same four-week period, then the estimated fare elasticity will be a biased estimate of the true elasticity.

This problem almost certainly does not affect our model. For the pre-privatisation period, fares were set annually by British Rail. Currently, fares are set in three general rounds: January, May and September, implying that train operators do not set prices in response to demand conditions in the same four-week period. The use of four-weekly data does however ensure that we fully take account of fares changes made at various points during the year.

4.5.3. Estimation of cross-price elasticities of demand

Although the primary purpose of our analysis is to derive robust estimates of the own-price elasticity of demand either for demand aggregated over all ticket types, or for the groupings of ticket types discussed in Section 4.4.2.2 above, the PDFH also reports cross-elasticities of demand between certain ticket types. For the reasons given above, we have not attempted to estimate demand elasticities for all of the ticket type groupings reported in the PDFH in respect of the two long distance panels. However, we have examined the cross-elasticities of demand between the ticket types in our disaggregated analysis, by including both the own price and the price of the other ticket type category in our demand equations.

The inclusion of both the own-price and the price of alternative ticket types in the demand equation raises a statistical problem when there is a strong correlation between the prices. This is the problem of multicollinearity and it results in a lack of accuracy in estimates the

[16] We have also considered the use of slope dummies. However, this would imply that the degree of price elasticity and therefore substitutability between rail and other modes had changed during and possibly after the period of severe post-Hatfield disruption. This seems much less likely than a simple shift in the level of demand, which can be explored using the type of Dummy variable we have used. Although people with good alternative means of making the journey might have switched out first, thus reducing the aggregate, those remaining would be closer to finding substitutes, so their elasticities would rise. Once the rail service returns to a more normal level, we should expect these patterns to reverse in any case. As we will see in Section 5.5.3, our pre- and post Hatfield elasticity estimates are indeed very similar for all but one of our panels. The exception is the London TCA and South East panel, where the difference is almost certainly caused by other factors (the Hatfield accident only had a minor impact on this segment).

effects that the explanatory variables produce on the dependent variable (i.e. elasticities of the model). If this data problem occurs, the regression analysis is not able to precisely allocate the impact of the explanatory variables on the dependent variable, since some or all of those explanatory variables are approximately similar to a linear combination of the other explanatory variables. In other words, if multicollinearity exists, the estimates of the coefficients may have the unexpected sign or an implausible magnitude.

We have investigated whether multicollinearity is likely to be a problem by examining the correlations between the prices of the ticket type groupings. Appendix D presents the results from this analysis. In conclusion, we find that there is evidence of collinearity between prices, and as such, the price elasticity estimates based on equations including both own price and a cross-price should be treated with caution.

In general, we believe that the estimation of cross-elasticities of demand using only econometric techniques is difficult. In time series models, variables are usually highly correlated and this makes it difficult to estimate cross elasticity due to the multicollinearity problems discussed above. We do not believe that proxy variables are able to solve the problem.[17] Rather, we believe that cross-elasticities should in general be estimated using survey work, the results of which can then be used to impose constraints on the cross-elasticities in the econometric model. This is the approach that has been used to estimate the current cross-elasticities in the PDFH, see Section 2.1.2 for a description.

[17] One possible approach might be the following. The pair of equations

$$D_1 = a_1 - b_1p_1 + c_1p_2 + u_1$$

$$D_2 = a_2 + b_2 p_1 - c_2p_2 + u_2$$

can be rewritten in terms of D_1, p_1, and D_2, for example:

$$D_1 = a_1 - b_1p_1 + (c_1/c_2)\{D_2 - a_2 - b_2 p_1 - u_2\} + u_1$$

If a means existed to get a value for c_2, then this could lead to an estimate of the cross elasticity c_1. However, the error term $u_1 - (c_1/c_2)u_2$ would in this case be correlated with one of the explanatory variables and this would lead to biased estimates. This could be compounded if u_1 and u_2 are correlated. Because of this, this is not a suitable approach.

4.6. Econometric Framework

4.6.1. Time series properties of data

The type of dataset we have assembled for this study is comprised of groups of pooled cross-section and time series data. The advantages of this type of dataset over simple cross-section or time-series data are the increases in the number of data points and in the degree of variability in the data; both of these help to increase the precision of the estimates. However, an important issue arises in the application of econometric modelling to time series data and panel data. Variables often tend to be highly correlated with their own past values and this creates a real danger that standard regression techniques may result in estimates which are not statistically robust. This is known as the problem of spurious regression and occurs when data are ***non-stationary***.

A data series is labelled non-stationary if there is no tendency for the series to revert to a constant mean. In most cases, this will be true when the data exhibit a growth trend. There are standard statistical tests available to determine whether or not time series are non-stationary. We have applied these tests to the panel dataset we have constructed and found that the demand, price and GDP data series are indeed all non-stationary.

Recent advances in econometric theory have led to the development of techniques that enable the researcher to estimate models using non-stationary data. To estimate such models consistently, the requirement is that there must be a stable long run relationship between the non-stationary variables in the model. This condition underlies the concept of co-integration.

Co-integration is more than simply an econometric issue. The concept of a long run fare elasticity of demand implicitly assumes that there exists a predictable long run proportional demand response to changes in fares. If there is no co-integrating relationship between the variables in our demand model, then the model contains no such predictability and as such, is not able to provide robust long-run elasticity estimates. In this instance, the model would certainly be mis-specified. However:

> If there is a cointegrating relationship between our explanatory variables, we are able to estimate <u>long run elasticities of demand</u> using equations specified in levels of variables.

In that case, our results would not be affected by spurious correlation.

We have employed recently developed econometric techniques in the field of panel co-integration analysis to verify that there is in fact a co-integrating demand relationship amongst the variables in our model. The tests that we have carried out are described in Appendix B.

4.6.2. Choice of econometric model

The primary objective of the present study is the estimation of long run fare elasticities of demand. This objective guides our choice of econometric model to the extent that we are willing to sacrifice overall explanatory power in terms of explaining the dynamics of demand responses in order to obtain increased precision of the estimates of the long run parameters.

We have established that there is a co-integrating relationship between the level of demand, prices and the level of GDP, and so we choose a model specified in levels of variables in order to estimate the long-run elasticity values. Further, since we are concerned with estimating fare elasticities for whole panels rather than for individual flows, we use a fixed effects panel framework, which constrains the elasticity parameters to be the same for each flow within panels. Since it is the elasticity for the panel as a whole that is of interest, the direct imposition of this restriction in the model leads to a reduction in the number of parameters that require estimation and hence an increase in the precision of the estimates.

The functional form of the fixed effects model is in a general form specified as follows:

$$ln\,D_{it} = b\;ln\,P_{it} + \boldsymbol{X}'_{it}\gamma + a_i + e_{it} \qquad \text{(ii)}$$

where:

$\boldsymbol{lnD_{it}}$	is the log of demand for rail travel on flow ***i*** in period ***t***;
$\boldsymbol{lnP_{it}}$	is the log of price of rail travel on flow ***i*** in period ***t***;
b	is the long-run fare elasticity for the panel of flows;
$\boldsymbol{X_{it}}$	is a vector of other explanatory variables for flow ***i*** in period ***t, such as GDP, prices of other tickets and other variables***[18]
γ	is a vector of panel long run elasticities w.r.t the explanatory variables;
a_i	is a fixed effect for flow ***i***
e_{it}	is a random disturbance

Since the explanatory variables are cointegrated, the model of demand presented above describes a long run relationship, and as such is appropriate for deriving long run fare elasticities. The double logarithmic specification entails that the elasticity is constant at all price levels and equal to the parameter b .

The model has two attractive features:

- First, few assumptions are required to ensure that it yields unbiased estimates. For example, in contrast to random effects panel models, there is no requirement that the

[18] See Section 4.6.3 for the explanatory variables that we have eventually used in each of our panels.

individual effects be uncorrelated with the explanatory variables, since the fixed effects are included themselves as regressors.

- Second, the model is efficient in the sense that the b parameter, which measures the long run fare elasticity, will have lower standard errors than the alternative Mean Group estimator, which requires the estimation of individual flow equations.

Dynamics of demand

The model represented by equation (ii) cannot provide any information on the dynamic adjustment process, ie the speed with which the effect of a price change converges to its long run effect on demand. We have investigated the dynamics of rail demand using a Vector Error Correction Model (see Box 4.1).

The Vector Error Correction Model (VECM) is wholly appropriate for the combined estimation of both long run estimates and short run dynamics in panel data. It is not, however, the most efficient method to estimate long run elasticities if these are, as in the present study, the main parameters of interest.

The VECM specification is derived from a straightforward re-parameterisation of the standard Auto-Regressive Distributed Lag (ARDL) model, in which a time-series variable is modelled as a function of its own lagged values and current and lagged values of its covariates. The long term fare elasticities for each ticket type within each panel are identified in the model by the average of the estimates for each flow within the same ticket type group. This is the Mean Group (MG) estimator, and it is a consistent estimator for the long run average elasticity of the group even where the parameters are heterogeneous across flows. In large samples, the estimates will therefore tend towards their true values.

Box 4.1
Vector Error Correction Model

The functional form of the Vector Error Correction Model that we adopt is specified as follows:

$$D\ln D_{it} = q_i(\ln D_{i,t-1} - b'_i\, \boldsymbol{X}_{i,t-1}) + \sum_{k=1}^{m-1} g_i D\ln D_{i,t-k} + \sum_{k=1}^{n-1} g_i' D\boldsymbol{X}_{i,t-k} + m_i + v_{it}$$

where b_i are the long-run parameters,
$\boldsymbol{X}_{it}$ are all the explanatory variables, measured in logs,
q_i are the error correction parameters,
γ_i are the parameters governing the dynamics of demand,
μ_i are the fixed effects, and
$\boldsymbol{v}_{it}$ are the random disturbances.

In this model, the expression in parenthesis is a measure of the divergence between actual demand and predicted demand given the long run elasticity parameters (b_i). The error correction parameters (q_i) measure the strength of the force pulling demand back towards its long run predicted level in the period after it has strayed away. The model is called an "Error Correction Model" because of this feature.

4.6.3. Final specification of demand models

The full set of explanatory variables described in Section 4.3 were tested for inclusion in the final models. The process involved testing variables to assess their contribution to explaining the variance of demand and discarding those found to be insignificant. We also tested different specifications of dummy variables to capture the effects of the Hatfield crash (see Section 4.5.1).

During this process of refinement, our measures of generalised journey time, service quality and the attributes of alternative transport modes were each found to be insignificant in all the demand models. The measure of economic activity was statistically significant only for the London TCA and South East panel of flows, and the equations for this panel include this variable in addition to the core variables - price, GDP, seasonal effects dummy variables and Hatfield dummy variables. Since we have literally run dozens of tests, all using different variables in different combinations with each other for each of the panels, we do not report in detail on the test results from this process of refinement. However, to give a flavour for the impacts that occurred we give the following examples for the London TCA to and from Rest of Country panel:

- the petrol price, unemployment and punctuality variables were all highly insignificant;
- including the vehicle kilometres variables rendered the GDP elasticity negative as the two variables are highly correlated; and
- including the generalised journey time variable resulted in implausibly low values for both price and GDP elasticity.[19]

Variables that were not significant were dropped from the model. This is consistent with established econometric practice which suggests that models should be run as economically as possible.

The final specification included two dummy variables in each equation to capture the effects of the Hatfield crash. The first captures the effect on demand in the three periods immediately following the crash and the second dummy variable captures the effect over the following three periods.

4.6.4. Estimation software used

We have carried out our analyses using the EViews 3.1 package. EViews 3.1 is a standard econometrics software package and was able to provide all the relevant data analysis and regression tools for this project. In particular, the software allowed estimating both the Fixed Effects Panel Model and the Error Correction Models. The Fixed Effects Panel Models were estimated by using the pooled estimation facilities of Eviews. For the Error Correction Models estimation, some specific programs in the EViews language have been created.

---

19 The difficulties with using GJT data to take account the impact of the Hatfield accident were discussed in Section 4.3.3. More generally, introducing a GJT variables may introduce endogeneity problems if GJT changes are caused by service level changes, which may themselves be a supply side response to an increase in demand.

5. RESULTS

5.1. Introduction

In this chapter, we present the results of our estimation work. We present and discuss summary results only; full results of the models for each of the panel (including t-statistics) may be found in Appendix C.

As noted in Section 4.5.3, the disaggregated analysis may have been affected by multicollinearity in our dataset. This has reduced the robustness of the disaggregate own-elasticity and caused particular problems in the estimation of cross-elasticities that are the subject of the present section.

Multicollinearity occurs when some of the explanatory variables are very highly correlated. In such a situation, it is difficult to disentangle the separate impact of each of the two variables on the dependent variable. This can lead to coefficients having an unexpected sign or magnitude, and very high standard errors attached to them.

We have tested for multicollinearity in the ticket type category variables and found evidence of potential multicollinearity in a number of panels. As a result, the disaggregate results need to be interpreted with care, in particular the cross elasticities that we will report in the various sections in this chapter. The full multicollinearity analysis is reported in Appendix D.

The structure of the chapter is as follows:

- Section 5.2 contains our estimates for the two long distance panels;
- Section 5.3 analyses the London and South East panel;
- Section 5.4 covers the Regional (Non London Short Distance panel); and
- Section 5.5 briefly discusses the impact of the Hatfield accident on rail demand.

5.2. Long Distance

5.2.1. Aggregate elasticities

Our aggregate price elasticities for the two long distance panels are reported in Table 5.1. The table includes both the long run elasticities estimated using the panel fixed effects estimation methods, and the short and long run elasticities estimated using the Vector Error Correction Model (VECM) estimation method (see Section 4.6.2).

Table 5.1
Aggregate Elasticity Estimates: Long Distance

	Panel fixed effects	VECM	
	Long term	Short term	Long term
London TCA to and from Rest of Country	-0.64	-0.20	-0.70
Non-London Long Distance	-0.82	-0.18	-1.01

Source: NERA estimates

The demand for travel in both long distance segments is inelastic with regard to price, with the elasticity in the non-London based segment being higher (in absolute terms) than in the flows to and from London.

The price elasticity for the London TCA to and from Rest of Country segment is (in absolute terms) well below unity in both the fixed effects and VECM models. The estimates resulting from these two models are also very close to each other. Although similar values were found by NERA in previous work for OPRAF, the values in the PDFH (between –0.9 and –1.0) suggest a more price-sensitive demand. Moreover, the PDFH recommendations are intended to represent the short-term response to demand, whereas the NERA values are long-term elasticities.

In the ***Non London Long Distance*** panel, where the PDFH recommended values are between –0.85 and –0.90, the NERA fixed effects and PDFH values are of the same order of magnitude, whereas the VECM estimate is somewhat higher.

The short-run VECM elasticities are based on the value of the adjustment factor, which is 0.277 for the London panel and 0.174 for the non-London panel.

5.2.2. Disaggregate elasticities

5.2.2.1. Own price elasticities

In the two long distance segments, the data have as noted in Section 4.4.2.2 been disaggregated into the following two ticket type categories:

- fully flexible tickets (both First and Standard Class); and
- reduced / restricted tickets.

Table 5.2 contains the estimated own price and GDP elasticities for the fully flexible tickets segment in the two long distance panels. The aggregate price elasticity values from Table 5.1 are also included for comparison purposes.

Table 5.2

Disaggregate Long-Term Elasticity Estimates: London TCA to and from Rest of Country

	Price	*(Aggregate price, fixed effects)*
Fully flexible tickets	-0.49	***-0.64***
Reduced / restricted tickets	-0.82	***-0.64***

Source: NERA estimates

The table shows that the demand for fully flexible tickets in this segment is clearly inelastic with regard to price. It is also more inelastic than aggregate demand in the panel. The PDFH recommends for London flows an own elasticity of between –0.60 and –0.90 (depending on distance) for First Class tickets and an own elasticity of between –1.00 and –1.50 for full Standard Class tickets. The NERA estimates suggest that the demand for premium tickets in the London-based panel may be substantially more inelastic.

By contrast, the demand for reduced / restricted tickets is more sensitive to fare levels than aggregate demand. It is however still inelastic with an estimated price elasticity of –0.82. This compares with PDFH recommendations for this segment of between –1.25 and –1.40, depending on distance.

Table 5.3 presents the disaggregated analysis for the Non London Long Distance panel. Here, the demand for reduced / restricted tickets is less sensitive to price than aggregate demand. The estimated price elasticity for reduced tickets in this segment is –0.70. By contrast, the demand for fully flexible segments in this segment is more elastic than aggregate demand. These disaggregated values cannot be compared with the PDFH recommendations since the PDFH does not contain disaggregated recommendations for non-London long distance flows.

Table 5.3

Disaggregate Long-Term Elasticity Estimates: Non London Long Distance

	Price	*(Aggregate price)*
Fully flexible tickets	-1.16	***-0.82***
Reduced / restricted tickets	-0.70	***-0.82***

Source: NERA estimates

We note that all disaggregated estimates may have been affected by multicollinearity problems in the data set. A detailed analysis of this issue is provided in Appendix D. The disaggregate values therefore need to be interpreted with care.

5.2.2.2. *Cross-price elasticities*

In the case of the two long distance panels, our disaggregated analysis has allowed us to derive cross-elasticity estimates with correct signs. We report them here but note that they may not be robust due to possible multicollinearity problems (see Appendix D).[20] As we have noted in Section 4.5.3, we believe that cross-elasticities should in general be estimated on the basis of survey work, the results of which can then be used as constraints on the cross-elasticities in the econometric model.

Our cross-elasticity estimates for the ***London TCA to and from Rest of Country*** panel are shown in Table 5.4. The interpretation of the values in the table is that if, for example, the price of restricted/reduced tickets rises by 10 per cent, and the price of fully flexible tickets remains unchanged, the demand for fully flexible tickets rises by 1.5 per cent. If on the other hand the price of fully flexible fares rises by 10 per cent, and the price of restricted/reduced tickets remains constant, the demand for reduced tickets rises by 2.6 per cent.[21]

Table 5.4
Long-Term Cross-Elasticity Estimates: London TCA to and from Rest of Country

	Fully flexible price	Reduced/restricted price
Demand for fully flexible tickets	-0.49	0.15
Demand for reduced/restricted tickets	0.26	-0.82

Source: NERA estimates

Table 5.5 contains our long-term cross elasticity estimates for the ***Non London Intercity*** segment. The results suggest that the substitutability of the two ticket type categories in this segment is significant. If reduced fares increase by 10 per cent, the demand for fully flexible tickets would increase by 7.4 per cent. The estimated cross elasticity of the demand for reduced tickets with respect to full fare levels is 0.48.

---

20 As can be seen in Appendix C.2, all cross-elasticities that we report here are significant in that the relevant coefficient have high t-values. However, in the presence of multicollinearity, a result being significant in terms of having a high t-value does not imply that it is also accurate. We do not therefore report the t-values in this section, since these would be misleading. Whether cross-elasticities in the presence of multicollinearity can be used is ultimately a matter of judgment and cannot be determined on the basis of a single “significance” figure.

21 Given that the reduced/restricted segment is much larger in volume terms than the fully flexible segment, one would expect the cross-elasticity of the demand for fully flexible tickets with respect to the price of reduced/restricted tickets to be larger than the cross-elasticity of the demand for reduced/restricted tickets with respect to the price of fully flexible tickets. In our estimates, the opposite is the case, suggesting they may not be robust.

Table 5.5
Long-Term Cross-Elasticity Estimates: Non London Intercity

	Fully flexible price	Reduced/restricted price
Demand for fully flexible tickets	-1.16	0.74
Demand for reduced / restricted tickets	0.48	-0.70

Source: NERA estimates

5.3. London and South East

5.3.1. Aggregate elasticities

Table 5.6 contains our aggregate long term elasticity estimates for our London and South East panel.

Table 5.6
Aggregate Elasticity Estimates: London TCA and South East

	Panel fixed effects	VECM	
	Long run	Short run	Long run
Both directions	-0.62	-0.16	-0.68
To London	***-0.81***		
From London	***-0.51***		

Source: NERA estimates

The estimated aggregate price elasticity is –0.62 using the panel fixed effects method and –0.68 using the VECM method. Although these estimates are close to each other, they are not entirely comparable. The panel fixed effects formulation includes the economic activity variable, which was highly significant and produced the best model fit. By contrast, the economic activity variable was not significant in the VECM model.[22] In addition, both estimates are sensitive to the time period used: see Section 5.5.3.

The PDFH recommendations for this segment lie between –0.3 and –0.72 (the PDFH does not provide aggregate values for this segment).

In line with the PDFH recommendations, we have also analysed travel to and from London separately. Of these, travel to London is of course by far the larger segment, as noted in

22 The panel fixed effects model without economic activity variables produced an estimated aggregate price elasticity of –0.82.

Section 4.4.2 above. In this market, we have found demand to be somewhat more sensitive to price than on aggregate, while the opposite applies to travel from London.

This result is contrary to the recommendations in the PDFH, which suggest that travel to London is more inelastic than travel from London. It is not easy to explain the finding. A closer examination of the data has suggested that while the trend in fares to London has been in line with the general picture presented in Section 3.3.2, the development in fares from London has been much more erratic. This is almost certainly due to the impact of fares regulation. In addition, the number of travellers to London is an order of magnitude greater than the number of travellers from London. We therefore regard the estimated elasticity value for travel to London as more robust than the one for travel from London.

The value of the adjustment factor, from which the short-term VECM elasticity values have been derived, is 0.233 for this panel.

5.3.2. Disaggregate elasticities

5.3.2.1. Own price elasticities

In regard to the London TCA to and from South East panel, our analysis has distinguished between the two following ticket type categories:

- season tickets; and
- non-season tickets.

Our estimated own elasticities for these two ticket type categories in this panel are shown in Table 5.7.

Table 5.7
Disaggregate Long-Term Elasticity Estimates: Short Distance, Season Tickets

	Price	*(Aggregate price)*
Season tickets[23]	-0.83	***-0.62***
Non-season tickets	-0.40	***-0.62***

Source: NERA estimates

The demand for season tickets in the South East is on the basis of our estimates more sensitive to price than aggregate demand, the demand for non-season tickets less. By

[23] We have also tried to break down the disaggregated analysis into to London/from London, but due to the quality of the data, this did not produce meaningful results.

contrast, the PDFH recommendations suggest that demand for season tickets is ***less*** sensitive to price than the demand for other tickets.

5.3.2.2. Cross-elasticities

Estimating cross-elasticities for this panel has been difficult. We have tried to explain the demand for non-season tickets in terms of the price of season and non-season tickets, the level of GDP and the level of economic activity. This however results in the price of non-season tickets no longer having a significant influence on the demand for non-season tickets. Instead, it is the price of ***season*** tickets that is significant but with a negative sign. Normally, however, cross-elasticities can be expected to be positive.

Because of these problems, we do not further report further on cross-elasticities for this panel.

5.4. Regional

5.4.1. Aggregate elasticities

Table 5.8 contains our estimated aggregate elasticity values for the Non London Short Distance panel on the basis of the two estimation methods that we have used. The long-term elasticity value estimated using the VECM estimation method is somewhat higher than that using the panel fixed effects estimation method. Although the PDFH values for this segment are given by ticket type and do not include aggregate values, they are generally below our estimated long-term values.

Table 5.8
Aggregate Elasticity Estimates: Non London Short Distance

	Panel fixed effects	VECM	
	Long run	Short run	Long run
Non-London Short Distance	-0.95	-0.34	-1.14

Source: NERA estimates

5.4.2. Disaggregate elasticities

5.4.2.1. Own-elasticities

Our own-elasticity estimates for this segment, disaggregated between season and non-season tickets, are shown in Table 5.9 below.

Table 5.9
Disaggregate Long-Term Elasticity Estimates: Non London Short Distance

	Price	*(Aggregate price)*
Season tickets	-1.50	***-0.95***
Non-season tickets	-1.21	***-0.95***

Source: NERA estimates

As can be seen, the demand for season tickets is considerably more elastic than aggregate demand. This however applies to non-season tickets as well (see below), which raises doubts as to the robustness of the disaggregation between ticket types in this panel. This may have been caused by multicollinearity problems: see Appendix D for a full analysis. The PDFH contains a uniform (short-term) season ticket price elasticity for this segment of –0.6.

5.4.2.2. Cross-elasticities

We have consistently found a negative cross-elasticity of the demand for season tickets with respect to the price of non-season tickets in this panel, which is contrary to expectations. Because of this, we do not report further on cross-elasticities in this segment.

5.5. The Impact of the Hatfield Accident

5.5.1. Introduction

While our research has focused on long-term elasticities, it has been necessary to take account of the impacts of the Hatfield accident and the subsequent disruption on demand. We have also examined the issue of the stability of the demand parameters by estimating equations both for the period prior to the Hatfield accident, and for the entire period for which data are available.

5.5.2. The effect of the post-Hatfield disruption on rail demand

Following some experimentation with different model specifications, including the use of data on train punctuality and reliability, we chose an approach to estimating the effect of the post-Hatfield disruptions based on the use of dummy variables in levels, which represent demand impacts in terms of a change in the constant term in the regression equation when the dummy variable is "on". We have specified two dummy variables in our equations, one covering the first three four-weekly periods (i.e. weeks 1 to 12) after the accident, the other one covering the second three (i.e. weeks 13 to 24). In this way, most of the impact of the accident should be captured by the dummy variables without impacting on our elasticity estimates.

The values of the dummy variables are interesting in their own right, since they give insight into both the magnitude of the demand shock after the accident, and also to some extent into how this varied by ticket type category.

The impact of the accident and subsequent disruptions to rail services on aggregate demand in each panel is shown in Table 5.10. It should be emphasised that the figures reported in the table only show the impact on demand that is directly due to the accident. Other impacts of demand, such as seasonal changes, changes in fares or GDP, have been isolated from these figures as they are accounted for by other variables in our equations.

Table 5.10
Change in Aggregate Demand Due to Hatfield Accident
(%)

	In week 1 to 12 after the accident	In week 13 to 24 after the accident
London TCA to and from Rest of Country	-24.7	-15.9
Non London Long Distance	-28.3	-15.2
London TCA and South East	not significant	not significant
Non London Short Distance	-10.9	-4.7

Source: NERA estimates

As expected, the most significant impact occurs on the long distance routes. On the short distance routes, a clear impact is also noticeable on the non-London routes, but not on the London routes. In all panels where the effect is significant, the impact in the period immediately following the accident, when service disruptions were most pronounced, is substantially greater than in the later period.

In addition, our analysis has also produced disaggregated impact values for two of the four panels. The panels for which we have reliable estimates are the London TCA to and from Rest of Country panel, and the London TCA and South East panel. In the other two panels, the disaggregated analysis did not produce meaningful results for the Hatfield variables.

The impacts of the Hatfield accident disaggregated by ticket type category are for the two London panels shown in Table 5.11.

Table 5.11
Change in Disaggregate Demand Due to Hatfield Accident

	In week 1 to 12 after the accident	In week 13 to 24 after the accident
London TCA to and from Rest of Country	-24.7	-15.9
Fully flexible	***-18.6***	***-8.5***
Reduced/restricted	***-27.2***	***-22.6***
London TCA and South East	not significant	not significant
Seasons	***+4.3***	***+5.9***
Non-seasons	***-6.8***	***not significant***

Source: NERA estimates

Although Table 5.11 shows interesting results, it needs to be interpreted with care. In regard to the London TCA to and from Rest of Country panel, the results suggest that the impact of the Hatfield accident was less severe in the fully flexible segment than in the reduced/restricted segment, suggesting, consistent with the results on fare elasticities reported above, that this segment of demand is somewhat more "captive" to rail than demand for restricted availability/advance purchase tickets. In the London TCA and South East panel, no impact is noticeable on aggregate demand. However, when disaggregating between ticket type categories, a negative impact does appear to exist in the non-seasons category. By contrast, the demand for season tickets appears to have increased following the accident. This is unlikely to reflect the impact of the accident itself but may be due to factors unexplained by our explanatory variables. In view of this, the results should be interpreted with caution.

5.5.3. Parameter stability

To test the robustness of our elasticity estimates, we have also estimated models for each panel in which only the period up to the Hatfield accident was analysed. Encouragingly, we found that in all but one case, the elasticity values for the pre-Hatfield period were very close to those estimated for the entire period.

The exception is the London TCA and South East panel. For this panel, we estimated an aggregate elasticity of –0.62 for the entire period. However, when only evaluating the period up to September 2000 (i.e. excluding the Hatfield accident), the estimated aggregate elasticity for this panel is –1.07. A similar result was obtained from the VECM model.

We do not believe that this difference is due to the Hatfield accident itself. The difference did not occur in the other panels and the Hatfield dummy variables should have picked up any impact. Moreover, this panel was least affected by the Hatfield accident anyway.

To examine this further, we have undertaken supplementary analysis using post-Hatfield data only. During this period, average revenue fell slightly, while GDP increased (but below trend). The results, which must be regarded as highly tentative because of the short time period, suggest that demand has over this period mainly been responsive to the changes in GDP and only to a limited extent to the slowly falling fares. The indicative price elasticity for this period is between –0.15 and –0.20, which explains the observed fall in the long run elasticity.

We are unable at this point to throw further light on the apparent discrepancy between the two estimates of price elasticity of demand in this panel. We comment further on the significance of this finding in interpreting our results in Chapter 6 below.

6. IMPLICATIONS FOR PDFH ELASTICITIES

6.1. Introduction

A handbook like the PDFH is used for a variety of purposes. For example, the introduction to the current edition of the PDFH notes that "its applications include investment appraisal, marketing and planning, budgeting and assessing the customer response to timetabling and operating decisions".

For some of these uses it would be most appropriate to have short-term elasticity values. Examples include short term budgeting decisions for the year ahead, assessing the customer response to timetabling decisions and short-term planning. For other types of uses, such as investment appraisal and longer-term strategic planning, long-term elasticities would be most helpful.

According to the PDFH, the currently recommended fare elasticity values in the PDFH are intended to represent "the change in annual demand after a change in that year's fares".[24] Chapter D9 of the PDFH also notes that the analysis underlying the PDFH recommendations is essentially static, with the emphasis on equilibrium states, and that the dynamic properties of rail demand are not currently considered in the PDFH. However, according to the PDFH, the limited amount of existing evidence suggests that demand for long established inter-city routes responds quickly to "marginal" changes in fares.

In the present chapter, we begin in Section 6.2 by examining the implications of our VECM (dynamic) estimation results for the interpretation of the elasticity values currently recommended in the PDFH. As we demonstrate in Section 6.2, the findings from the VECM estimations indicate that the PDFH elasticities may, in fact be almost equivalent to long-term elasticities. In Section 6.3, we examine whether our findings on long run elasticities support the current PDFH recommended values. In Section 6.4, we indicate how the findings from our VECM estimations can be used to provide additional information on short term demand responses, for use in short term business and financial planning. We provide a summary of our recommendations in Section 6.5.

6.2. Interpreting the PDFH Elasticities

In Chapter 5, we have reported the short and long term elasticities estimated using the Vector Error Correction Model (VECM). One of the attractive properties of this model is that it can also be used to estimate the speed with which the demand for rail travel adjusts to its long-term values. An overview of the insights arising from this analysis is provided in Table 6.1.

---

24 See PDFH, Chapter B2, page 23.

Table 6.1
Speed of Adjustment of Demand

	London TCA to and from Rest of Country	Non-London Long Distance	London TCA and South East	Non-London Short Distance
Value of the adjustment factor	0.277	0.174	0.233	0.298
Proportion of demand change occurring within four weeks	28%	17%	23%	30%
Proportion of demand change occurring within six months	88%	71%	82%	88%
Proportion of demand change occurring within one year	99%	92%	97%	99%
Amount of time required for the change in demand to adjust to within 1 per cent of its long-term value	1 yr 1 month	1 yr 10 months	1 yr 5 months	1 yr

Source: NERA estimates

As can be seen in the table, our estimates suggest that elasticities of demand in the rail sector approach their long run values over a period of between one and two years, depending on the type of flow. The table also suggests that for all panels, ***more than 90 per cent of the demand adjustment occurs within a year after the fares change.*** Even in the case of the Non London Long Distance panel, where it takes almost two years for demand to adjust to its long run level, about 92 per cent of this adjustment occurs within the first year.

This has important implications for the interpretation of the PDFH recommended elasticity values. If, as the PDFH invites us to, we interpret these values as the proportionate change in demand caused by a 1 per cent change in real fares, assessed one year after the fare change has occurred, it is clear from the results shown in Table 6.1 that the PDFH elasticities may be very close to being long run elasticities, in the sense that they capture between 92 per cent and 99 per cent of the long run demand response.

6.3. Long Run Elasticities

On the assumption that the existing PDFH recommended values can be interpreted as long run elasticities, we now consider the consistency between these values, and the estimates of long run elasticities presented in Chapter 5 above. The discussion is organised according to market segment as follows:-

- Section 6.3.1; London TCA to and from Rest of Country;
- Section 6.3.2; Non-London Long Distance;
- Section 6.3.3; London TCA and South East; and
- Section 6.3.4; Non-London Short Distance.

6.3.1. London TCA to and from Rest of Country

6.3.1.1. Aggregate elasticities

In this panel, NERA has estimated an aggregate fare elasticity of –0.64. Our results are consistent with previous NERA work for OPRAF, in which we estimated an aggregate elasticity for this segment of between –0.61 and –0.78. By contrast, the PDFH recommends aggregate elasticities of between –0.90 and -1.00, depending on distance. Because our VECM estimates indicate that virtually all of the long run effect occurs within one year for this panel, the PDFH elasticities can be regarded as long run elasticities.

The difference between the NERA and PDFH values is striking, in particular because the PDFH values for this segment are also based on recent econometric work (the Handbook Update study). Unfortunately, the econometric work underlying the PDFH remains undocumented and we have therefore been unable to examine the reasons for the differences in the estimates of fare elasticity.

Given the importance of this fare elasticity value for policy purposes, for example fare regulation, we recommend that some further work be undertaken in order to establish how the results from the NERA and the Handbook Update study can be reconciled with each other. A relatively brief review of econometric models and variables used could provide important insights. On the basis of the results of the additional work, a decision can then be made as to whether the PDFH elasticity values should be adjusted downwards.

6.3.1.2. Disaggregate elasticities

NERA has estimated an own-elasticity of fully flexible tickets of –0.49, and an own-elasticity of reduced/restricted tickets of –0.82. While the PDFH elasticities are higher in aggregate, the relative magnitude of the disaggregate estimates is not dissimilar to that of NERA's estimates. For that reason, we see no reason to adjust the relative values of the disaggregate elasticity values of the PDFH, although all disaggregate values are of course affected by the discrepancy in aggregate values as discussed above.

6.3.1.3. Cross-elasticities

While NERA has derived positive cross-elasticities between the two ticket type categories in this segment, we do not regard the results as very robust. We have drawn attention in Chapter 5 above to the econometric problems stemming from the high degree of multicollinearity in the fares data series. We have also noted problems with the quality of the data set, in spite of extensive data cleaning, and it is also likely that the results are influenced by changes in ticket conditions and availability. We regard the approach taken in the Handbook Update study, where cross-elasticities were effectively imposed on the dataset using diversion factor relationships estimated from stated preference research, as preferable. For this reason, we do not recommend any change to the cross-elasticities in the PDFH.

6.3.1.4. Extent of disaggregation in PDFH elasticities

The PDFH elasticity values for the London TCA to and from Rest of Country segment are highly disaggregated. Values are broken down by distance band, ticket type category and fare per mile paid, resulting in dozens of possible combinations.

Given the serious problems that we have had in deriving robust estimates even at modest levels of disaggregation, we have strong doubts as to the robustness of these recommendations. In addition, they can sometimes be misleading. For example, the modifiers according to the fare per mile paid appear to be based on the presumption that elasticities will (in absolute terms) become higher if fares on a particular route are increased towards the revenue-maximising level. While such an effect can indeed be expected if other factors are held constant, it completely ignores the effects of other factors influencing fares per mile, in particular service quality, which will affect marginal costs, and the strength of competition from other modes, which will affect the profit-maximising mark-up over marginal cost.

We would therefore recommend simplifying the recommended elasticities by removing the fares per mile adjustment, and possibly merging the "first" and "full" categories into a single "fully flexible" segment, which is the approach taken in the present study. Instead of the mechanical fares per mile adjustment, qualitative guidance should be given as to how local factors may impact on elasticity values (this would be an extension of the current guidance on competition in Chapter B2.11 of the Handbook).

6.3.2. Non- London long distance flows

6.3.2.1. Aggregate elasticities

NERA's estimated long-term aggregate elasticity for the Non London Long Distance segment is –0.82, compared to the PDFH recommended values of between –0.85 and –0.90. We estimate that around 92 per cent of the long run effect occurs within a year for this panel, so that the PDFH values would imply long run elasticities of around -0.92 and –0.98.

In Section 6.3.1.1 above, we have recommended some comparison work between the NERA work and the Handbook Update study in order to explain the differences between the NERA and PDFH values in the London long distance segment. If it were to be undertaken, this work would probably also shed light on the differences between the estimates in the non-London long distance flows. However the magnitude of the difference is much smaller here, and our findings can be interpreted as being broadly consistent with the PDFH recommended values.

6.3.2.2. Disaggregate elasticities

For this segment, the PDFH does not provide recommended values by ticket type category. NERA has distinguished between fully flexible on the one hand and reduced/restricted on the other hand and found significant differences: a value of –1.16 for fully flexible tickets and of –0.70 for reduced / restricted tickets.

Because of the substantial difference between these two estimates, we recommend that a distinction between these ticket type categories be introduced in the PDFH. However, we note that the relationship between the elasticities for the two types of ticket is quite different to that estimated in the London-based long distance panel, and we are not sufficiently confident in our results, again because of data problems, to propose them as the basis for PDFH recommended values.

6.3.2.3. Cross-elasticities

NERA has derived positive cross-elasticities for this segment but we do not regard their magnitude as plausible. If cross-elasticities between the two ticket type categories were to be required, further work would be needed, probably using the diversion factor approach.

6.3.2.4. Extent of disaggregation in PDFH elasticities

For the reasons given in Section 6.3.1.4 above, we recommend replacing the current disaggregation according to the fare per mile paid by qualitative guidance on variations in elasticity values according to local factors. In addition, we suggest removing the present distinction between "Non London Long Distance (without full set of tickets)" and "Non London Long Distance (with full set of tickets)". The difference between the aggregate recommended values in each category (-0.85 versus –0.90) is in our view too small to warrant separate categories.

6.3.3. London TCA and South East

6.3.3.1. Aggregate elasticities

The aggregate price elasticity estimated by NERA for this panel is –0.62. The PDFH does not provide an aggregate value for this segment. However, its recommendations for disaggregated elasticities vary between –0.3 and –0.72. We therefore discuss the values in the next subsection. We note here that with the exception of the –0.3 value in the PDFH (for season tickets to London), the magnitude of the aggregate NERA and PDFH values is similar.

In Section 5.5.3, we have seen that the NERA estimate is sensitive to the time period used: when excluding the period since October 2000 from the dataset, a higher elasticity (in absolute terms) is obtained. Because of this, the estimate must be interpreted with caution.

6.3.3.2. Disaggregate elasticities

A comparison between the disaggregate PDFH and NERA values for this segment is provided in Table 6.2.

Table 6.2
Comparison of PDFH and NERA Elasticities for London TCA and South East Segment

	To London		From London	
	Seasons	Other	Seasons	Other
PDFH	-0.3	-0.66	-0.6	-0.72
NERA	-0.81		-0.51	
NERA seasons		-0.83		
NERA non-seasons		-0.55		
NERA aggregate		-0.62		

Source: PDFH; NERA estimates

The following two points should be noted about the results reported in the table:

- In the PDFH, elasticities for travel ***to*** London are lower (in absolute terms) than those for travel ***from*** London. In the NERA estimates, the picture is reversed.
- In the PDFH, elasticities for season ticket travel are lower than those for travel on other tickets. Here too, the NERA estimates suggest the opposite.

In regard to the distinction between travel to and from London, our prior expectation would be that car travel would be a closer substitute for rail travel on journeys from London for users with cars available, because of the greater availability and lower cost of car parking at the non-London end of the trip. On the other hand, car ownership is known to be significantly higher outside than inside London, so the proportion of rail users making outward journeys with cars available might well be smaller than the proportion of rail users travelling into London with cars available. It is not immediately clear how these factors would balance out. However, the distinction between inward and outward travel may be of little commercial relevance, given that ticket prices are not, and cannot be, differentiated by direction of travel, and also that the travel ***from*** London segment is relatively small compared to travel into London, and we would recommend the use of a single demand elasticity, covering both inbound and outbound journeys in PDFH..

The more interesting finding is the higher elasticity for season tickets travel than for travel on other tickets, which contrasts strongly with the relationship between the recommended values of the two types of ticket in PDFH. In making comparisons between the results obtained in the present study and the PDFH recommendations, it is worth bearing in mind that the PDFH material is not based on findings from recent research. It appears that the recommendations for South East flows are largely derived from a review of price sensitivity carried out by Steer Davies Gleave in 1993. The work involved a review of all previous work, and was complemented by a major stated preference study. However, it can be argued that there have been significant changes in market conditions in the South East over the past 20 years; we have referred above to the possible effects of changes in working patterns on the demand for different types of ticket. It could also be argued that demographic and employment location shifts, such as the growth in employment in satellite offices located around the M25, have altered employment opportunities for many clerical and administrative employees who would otherwise have travelled into Central London.

Whilst we do not have sufficient confidence in our results to argue that they should form the basis for PDFH recommendations, we believe that they do justify a more detailed re-appraisal of the PDFH parameter values.

If the long term elasticity for season tickets is higher than for other tickets, and if the market for travel ***to*** London is to a stronger extent dominated by season tickets than the market ***from*** London, then this might also form an explanation for why long-term elasticities ***to*** London may be higher than ***from*** London.[25]

6.3.3.3. Cross-elasticities

NERA has not been able to derive reliable cross-elasticities for the London TCA and South East panel. Our recommendation is for the PDFH to continue not to report cross-elasticities for this segment.

6.3.3.4. Extent of disaggregation in PDFH elasticities

We believe that the current extent of disaggregation in PDFH elasticities for the London TCA and South East panel should be retained.

6.3.4. Non London Short Distance

6.3.4.1. Aggregate elasticities

For the Non London Short Distance segment, NERA has estimated an aggregate long-term elasticity of –0.95. The PDFH values are –0.6 for season tickets, -0.9 for prepaid off-peak tickets and between -0.35 and –1.00 for daily tickets.[26] We would interpret our results as being broadly consistent with the PDFH recommended values in indicating that demand is significantly more price elastic in these markets than the demand for short distance travel into and out of London. This conforms with a priori expectations based on our perception of difference in the strength of competition between rail and other modes of transport between London and other major urban centres.

[25] We have tried to produce elasticity estimates disaggregated both by direction (to/from London) and by ticket type (seasons/non seasons), but due to the limitations of our data this unfortunately did not produce any meaningful results.

[26] Recommended values for daily tickets are broken down by PTE, Cross Boundary and Non PTE. For season and prepaid off-peak tickets, recommended values are identical across these area types.

6.3.4.2. Disaggregate elasticities

Unfortunately, due to significant data problems in this panel (in part caused by changes in ticket conditions over the study period, e.g. introduction and withdrawals of day return tickets), we have not been able to obtain any meaningful disaggregate elasticities for this panel. We deal with the appropriate extent of disaggregation in Section 6.3.4.4 below.

6.3.4.3. Cross-elasticities

NERA has not been able to derive reliable cross-elasticities for the Non London Short Distance panel. Our recommendation is for the PDFH to continue not to report cross-elasticities for this segment.

6.3.4.4. Extent of disaggregation in PDFH elasticities

We have some doubts as to the extent of disaggregation of the currently recommended PDFH elasticities for this panel.

Our first concern relates to the elasticities for daily tickets broken down by fare per mile. For reasons given in Section 6.3.1.4 above, we are sceptical as to whether such a differentiation is justified. In addition, providing such a disaggregation for some ticket types but not for others is inconsistent. For example, if elasticities for daily tickets in PTE areas do indeed vary with a factor 2.5 between low fare (£0.04 per mile) and high fare (£0.12 per mile) routes, then it is difficult to see why such a variation would not also exist for season and prepaid off-peak tickets.

In principle, we regard the distinction between PTE, Cross Boundary and Non PTE in this segment as potentially justified in view of the heterogeneity of the Non London Short Distance segment. However, it is again puzzling why there is such a large difference in the recommended elasticity values for daily tickets across the three categories, but no difference at all for season and pre-paid off-peak tickets.

Our recommendation would be to simplify the recommended values in this panel around the estimated NERA long term aggregate elasticity of –0.95. A disaggregation by ticket type and by area type would be desirable, but is difficult to achieve with econometric techniques, and may instead require further research using stated preference techniques. We would not recommend providing a quantitative breakdown according to the fare per mile paid. Instead, we would suggest providing some qualitative guidance on possible local factors that may influence the elasticity estimates, similar to the guidance on the impacts of competition currently provided in Chapter B2.11 of the Handbook.

6.4. Short Run Demand Response

If the elasticity values recommended in PDFH are to be interpreted as long run elasticities, the VECM estimation results reported in Chapter 5 provide a starting point for the type of guidance on short run demand response that might be offered in PDFH.

Figure 6.1 shows the speed of adjustment in aggregate demand in each of the four panels, with the number of elapsed four-weekly periods since the fare change on the horizontal axis and the long-term change in demand normalised to one. The figure indicates, consistent with the results shown in Table 6.1 above, that the speed of adjustment varies between the market segments, with a relatively rapid adjustment in the London based Inter City panel, and a slower adjustment in the non-London InterCity panel.

Figure 6.1
Speed of Adjustment in Demand by Market Segment
(assuming the long-term change in demand is 1 per cent)

Source: NERA estimates

Data of this kind, which would allow business managers in the TOCs to estimate the effect of fare changes on revenue within the twelve month period following a fares change, could be used as input to short term financial planning and budgeting procedures. The adjustment function for the non-London InterCity panel, for example, would suggest that the patronage loss in the first year after a price change might be approximately a third lower than the long run patronage impact, resulting in a larger revenue gain in the first year than in the long run.

We have only attempted to estimate reaction functions of the kind illustrated in Figure 6.1 for aggregate demand in each of the panels. At this stage, we would regard our results as indicative, and would recommend that the issue of short run adjustment using the type of dynamic model we have estimated merits further research. However, this approach appears to offer a means of addressing what is a significant lacuna in the present PDFH recommendations.

6.5. Summary and Recommendations

In this Chapter, we have reviewed the implications of our results for the recommended elasticity values in the PDFH. An important insight from the analysis in this chapter is that rail demand generally adjusts quickly to its long-run value, with the large majority of the demand impact occurring within one year of a shock. For this reason, the PDFH values, which represent demand response one year after a fares change, can plausibly be interpreted as long-term values.

We have also reviewed the estimates that we have obtained and discussed the implications of these. We have seen that some of our estimates are robust, other less so.

We regard as robust the estimates for our aggregate elasticities for the four panels. We generally obtained similar aggregate results from our panel fixed effects and VECM models. However, where our aggregate results differ significantly from the PDFH recommended values, we first recommend some further research to see whether explanations for the differences in values can be found. This applies in particular to the London TCA to and from Rest of Country panel.

We regard as non-robust our_estimates of most disaggregated and cross elasticities. While we have made a serious attempt to estimate these elasticities, in the end we are not sufficiently confident in them to recommend them as basis for changes to the PDFH values. One of the reasons for this is that too many changes in definition and availability of ticket types occur over time to render the CAPRI database suitable for disaggregated time-series analyses. This applies even in spite of the very extensive data cleaning exercise that we have undertaken. Another reason is that fare changes of different ticket types are highly correlated with each other, giving rise to potential multicollinearity problems.

We do however believe that in some cases, our disaggregated values suggest a case for a reappraisal of the values in PDFH. In general, we believe that cross-elasticities can best be estimated by analysis based on stated rather than revealed preference data, the results of which can be then used to impose constraints on the cross-elasticities in a time series econometric model.

Our key recommendations for further research are as follows:

- a comparison of the methodology and results used in the present study and the study from which the recommendations for the London TCA to and from Rest of Country segment were derived (see Section 6.3.1.1);
- research into cross-elasticities in the non-London long distance panel, to support the introduction of recommended values in the PDFH (see Section 6.3.2.2);
- a reappraisal of the disaggregated recommended values for the London TCA and South East segment (see Section 6.3.3.2); and
- research into a better disaggregation of elasticity values in the Non London Short Distance segment by ticket type category and area type (see Section 6.3.4.4).

7. DEVELOPMENT OF AN UPDATING MECHANISM

7.1. Introduction

After our discussion of the implications of our newly derived long-term elasticities for the PDFH in Chapter 6, we present in the current chapter a broad outline of a methodology that SRA could use to update elasticities on a regular basis.

The present version of the PDFH (version 4) dates from August 2002. The history of the PDFH is shown in Table 7.1. The Table shows that new versions of the PDFH have been published, on average, every four years.

Table 7.1
History of the Passenger Demand Forecasting Handbook

Issue	Year
British Rail Passenger Forecasts	1982
PDFH version 1	1986
PDFH version 2	1989
PDFH version 3	1994 / 1997
PDFH version 4	2002

A major reason for updating the PDFH elasticities from time to time arises from the fact that internal and external factors influencing the sensitivity of rail demand to fare adjustments change over time. Examples of such changes are the increased congestion and higher real fuel prices that were seen in the 1990s, compared to earlier periods, which might have led to lower fare elasticities (in absolute terms).

For any given set of elasticities there should therefore be no need to update the elasticity values unless there are significant changes in the factors influencing the sensitivity of rail demand to fare changes. However, changes to the rail business may mean that the elasticities that are relevant may change over time.

In our view, an updating mechanism should therefore consist of a three-stage process:

- Stage 1: Reviewing the purposes for which elasticity estimates are required. Are the current estimates appropriate in principle to the demands likely to be made on them?
- Stage 2: If they remain appropriate, deciding whether at any point in time an update of the elasticities should in fact take place.
- Stage 3: The derivation of the new elasticity values.

We will discuss each stage in turn in the next three sections.

7.2. Reviewing Appropriateness

The context in which PDFH estimates are used has changed greatly since privatisation. In the context of a single rail organisation, the same estimates could be used not only for testing the impact of a change in the tariff level or structure on demand and revenue, but also for evaluating investment projects.

Given the wide variety of uses to which estimates could be put, the term ***forecasting*** should perhaps be broadened to include the ***calibration of quantitative models*** used for assessing changes.

Under the current regime decisions about prices and investment are made in a completely different way. For some journeys fares are regulated, others can be varied at the discretion of the TOCs. The latter need to take account of both inter-modal and in some cases on-rail competition. Where there is on-rail competition the degree of aggregation implicit in both the CAPRI data set and the PDFH is not necessarily appropriate.

A good illustration of this is the fact of the decline in market share of the ticket types defined during the BR era, and the rise in “Other” (type 14) tickets. As discussed earlier, these appear to be a residual category incorporating special fares and fares restricted to single operators. Such a change creates challenges for data collection, demand estimation, and the interpretation of econometric results, as we discuss elsewhere in this report.

7.3. Deciding Whether an Update Should Take Place

Abstracting from these broader issues, one of the main purposes of the elasticities reported in the PDFH, as the name of the Handbook indicates, is demand forecasting. But if underlying elasticity values have changed, forecasts produced with the values in the Handbook will no longer be accurate. It follows that one way to test the elasticity values in the PDFH is by assessing how well they have performed against outturn data. The performance of the existing model should be the decisive factor in determining whether an update should take place.

In evaluating the forecasting performance of the current elasticities, it is important not to go into too much detail. For a number of market segments, for example London TCA to and from Rest of Country, the PDFH provides very detailed elasticity values that are broken down not only by ticket type category but also by fare per mile paid. Even when aggregating the PDFH elasticities to the maximum extent possible, a total of 18 elasticity values remain. These are shown in Table 7.2 below.

Table 7.2
Aggregate Elasticity Values by Market Segment in PDFH

			Commuting	Business	Leisure
London Travelcard Area			-0.3	-0.2	-1.0
		To London		From London/Non London	
		Seasons	***Other***	***Seasons***	***Other***
London TCA and South East		-0.3	-0.66	-0.6	-0.72
			20 to 100 miles	101 to 200 miles	Over 200 miles
London TCA to/from Rest of Country			-0.9	-1.0	-1.0
		With full set of tickets		Without full set of tickets	
		-0.90			
Non-London Long Distance			-0.85		-0.90
	Seasons	Prepaid Off-peak		Daily	
			PTE	***Cross Boundary***	***Non PTE***
Non-London Short Distance	-0.6	-0.9	-0.35	-1.00	-0.75
					Overall
To and From Airports					-0.5

Source: Passenger Demand Forecasting Handbook, August 2002

With the exception of the elasticities for the London Travelcard area,[27] the forecasting accuracy of all elasticities shown in Table 7.2 can be tested using CAPRI data (or other suitable data).

The tests should take account of the other key factors influencing rail demand, such as the level of real GDP.

No model is completely accurate, but each model will have a forecast error associated with it. The normal test for parameter stability is based upon comparing the actual error of the forecast with the error to be expected in the light of the goodness of fit of the data to the previous model. There are standard metrics for evaluating forecast outturns in comparison with, for example, the standard error of the forecast. This can for example be done using the EViews software packages, which we have also used for the present study. An example of the functionality that EViews provides in this area is given in Appendix E. In principle this allows an update advisory message to be triggered when the forecast error falls outside a pre-defined region.

The PDFH does not contain models as such, just elasticity estimates, which do not by themselves constitute a forecasting tool. Different models constructed with different ends might have different short-term and long-term forecasting performances.

[27] The values for the London Travelcard area cannot easily be tested since they are broken down by journey purpose as opposed to ticket type. These should therefore not be part of the updating mechanism that we set out below; instead it would be sensible to update these separately in collaboration with Transport for London.

This means that a key stage in the evaluation of the need for updating is to have an agreed model whose performance will provide the criterion for model updating. We recommend that the assessment of the performance of existing models should be undertaken on an annual basis. The best way of undertaking the assessment would be to develop a model that contains the relevant elasticity values and is able to undertake the assessment once outturn data are input into it. The model could be constructed using a professional econometric software packages like EViews (see the example in Appendix E), or alternatively could be in spreadsheet form which will allow it also to be used by those not familiar with professional econometrics packages. The model only needs to be developed once; in future years it can simply be run by adding another line of data.[28]

In a number of cases elasticities have been derived from stated preference surveys - for example, hypothetical switching behaviour. Some of these may not be amenable to direct econometric testing or validation. Nevertheless, the principle remains that updates are required in the light of one of two kinds of evidence:

- a priori evidence that the structure of the market and product offerings has changed; or
- ex post evidence that the model is performing poorly in the context in which it is being applied.

7.4. Undertaking the Update

If, on the basis of the assessment procedure described above, a decision has been made to update the PDFH elasticities, it is possible to do this in two different ways. These are:

- by refining the existing models; or
- by developing a completely new model.

In most situations it will be sufficient to update existing models. The exceptions occur:

- where there are fundamental changes in the variables determining the demand for rail travel. An example in the context of external factors influencing rail demand may have been the slowdown in the increase in car ownership that was seen during the 1990s;
- where the range of “products” on offer changes fundamentally so that existing classifications are no longer appropriate, at least for disaggregated analysis; and

---

[28] We note that NERA has in the past developed a spreadsheet-based forecasting framework for SRA’s predecessor, OPRAF.

- where new estimation methods become available that are clearly superior to tried and tested approaches. In the research underpinning this report we have explored a number of alternative approaches and have given some indication of their properties.

In the majority of cases, therefore, an update will take place by refining existing models, particularly for the more useful and meaningful aggregated analysis. Updating existing models will involve a number of standard steps. It is therefore possible to give general guidance on how existing models should be developed. We recommend that this be done by producing a Handbook on Updating PDFH Elasticity Models.

The Handbook should set out the steps that need to be followed. These steps bear some resemblance to the process that NERA has followed in producing our own estimates, which we have described in Chapter 4. We believe the key steps are the following:

1. Update the data bank. This involves checking for consistency with previous versions of the data, looking at data definitions, aggregation, and discrepancies in the actual numbers.

2. Identify problems with the data. Are they a credible representation of what they claim to measure? If not, can we track down sources of data error or inconsistency in definitions? The size of the problem will depend on the availability of informed people with a stake in ensuring data accuracy and consistency with established definitions.

3. Clean the data. There may be new information which will improve the quality of the old data, but the new data are most likely to have the biggest problems. Where obvious errors cannot be corrected, but only one or two observations are affected, it may sometimes be preferable to use reasonable interpolations rather than simply omit data, especially in short-term forecasting models where the presence of lagged values means that one missing value leads to the dropping of several data points.

4. Check the forecasting performance of the existing model on the new data. In a sense this will already have been done, but not necessarily at the detailed level or with corrected data.

5. Identify new causal factors that need to be incorporated in the model. (In the present study, the impact of Hatfield is an obvious example.) Where possible collect appropriate new variables for possible inclusion in the model.

6. Identify known issues relating to the estimation methods used. Once we get beyond very simple models, no estimation method is ideal, and choice is often a matter of a trade off between several different criteria, such as small and large sample unbiasedness, efficiency, simplicity and ease of understanding, and sensitivity to small variations in data. While simpler models tend to have virtues of robustness to new information and ease of understanding, they may not be the most efficient or with the smallest bias.

7. Adopt a core estimation approach, but carry out sensitivity analysis by trials of alternative, perhaps theoretically superior estimation methods. The core approach needs to be robust, without showing signs of substantial or systematic bias relative to more sophisticated methods.

8. Explore whether here has been a structural break in the model using standard techniques such as Chow tests. Alternatively, the inclusion of new variables reflecting new causal factors may mean that a single model fits the entire period over which data are available. In the event of a structural break estimation may need to be based on the post-break period. Usually one would look for a cause outside the included variables. For example, we have in the past examined the privatisation period as marking a change in the demand relationship.

9. Once the estimates have been derived check on the goodness of fit of the data and the consistency of the estimates with extraneous information. If elasticities have changed have they moved in the direction and of the order of magnitude that events would lead one to expect? Do we now understand why the previous model's forecasts became less reliable?

10. Incorporate the new estimates into forecasting software or other modelling applications that use the elasticities.

11. Ensure that appropriate data for future updates continues to be collected in a suitable form with adequate safeguards as to its accuracy. We would recommend that the data be collected in the same form as the database that NERA has constructed for the purpose of the present study. A summary of the structure of the dataset is provided in Table 7.3.

Table 7.3
Structure of NERA Dataset

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="4">Panels</th></tr>
<tr><th>London TCA to and from Rest of Country</th><th>London TCA and South East</th><th>Non-London Long Distance</th><th>Non-London Short Distance</th></tr>
<tr><td rowspan="14">Ticket types</td><td>1ST/EXEC</td><td rowspan="3">First+Standard</td><td rowspan="8">Other</td><td rowspan="3">First+Standard</td><td rowspan="8">Other</td></tr>
<tr><td>STD SINGLE</td></tr>
<tr><td>STD/OPEN RTN</td></tr>
<tr><td>SAVER</td><td rowspan="11">Reduced+Other</td><td rowspan="11">Reduced+Other</td></tr>
<tr><td>SUPERSAVER</td></tr>
<tr><td>NETWWK AWYBRK</td></tr>
<tr><td>APEX</td></tr>
<tr><td>DAY RETURN</td></tr>
<tr><td>ANNUAL S/T</td><td rowspan="3">Seasons</td><td rowspan="3">Seasons</td></tr>
<tr><td>WEEKLY S/T</td></tr>
<tr><td>OTHER SEASON</td></tr>
<tr><td>SLEEPER SUPP</td><td rowspan="3">Other</td><td rowspan="3">Other</td></tr>
<tr><td>CHDY SINGLE</td></tr>
<tr><td>OTHER</td></tr>
</table>

APPENDIX A. FLOWS CONTAINED IN EACH PANEL

Table A.1
Origin-Destination Pairs: London TCA to and from the Rest of Country

Origin Zone		Destination Zone	
000	London	572	Bath
000	London	300	Birmingham
000	London	570	Bristol
000	London	550	Cardiff
000	London	170	Carlisle
000	London	950	Edinburgh
000	London	960	Glasgow
000	London	230	Leeds
000	London	400	Leicester
000	London	130	Liverpool
000	London	100	Manchester
000	London	200	Newcastle
000	London	460	Norwich
000	London	410	Nottingham
000	London	510	Plymouth
000	London	150	Preston
000	London	103	Stockport
000	London	540	Swansea
000	London	580	Swindon
000	London	220	York

Table A.2
Origin-Destination Pairs: Non London Long Distance

Origin Zone		Destination Zone	
57	Avon	30	West Mids inner
57	Avon	42	Derby
57	Avon	13	South Merseyside
57	Avon	10	South Manchester
57	Avon	29	South Yorkshire
23	Leeds area	30	West Mids inner
23	Leeds area	42	Derby
23	Leeds area	13	South Merseyside
23	Leeds area	10	South Manchester
23	Leeds area	29	South Yorkshire
40	East Mids south	30	West Mids inner
40	East Mids south	42	Derby
40	East Mids south	13	South Merseyside
40	East Mids south	10	South Manchester
40	East Mids south	29	South Yorkshire
20	Tyne & Wear	42	Derby
20	Tyne & Wear	13	South Merseyside
20	Tyne & Wear	10	South Manchester
20	Tyne & Wear	29	South Yorkshire
41	Nottingham	30	West Mids inner
41	Nottingham	42	Derby
41	Nottingham	13	South Merseyside
41	Nottingham	10	South Manchester
41	Nottingham	29	South Yorkshire
45	Peterborough	47	Ipswich
30	West Mids inner	45	Peterborough
23	Leeds area	95	Edinburgh area
54	South Wales	57	Avon
17	Carlisle	34	Stoke
17	Carlisle	20	Tyne & Wear
23	Leeds area	27	North Humber
95	Edinburgh area	96	Glasgow inner
38	Rugby	96	Glasgow inner
21	Teesside	51	West Devon

Table A.3
Origin-Destination Pairs: London TCA to and South East

Origin Zone		**Destination Zone**	
000	London	01	Kent Link
000	London	02	Croydon Lines
000	London	03	Sutton Lines
000	London	04	S.W.Suburban
000	London	05	W.London / LHR
000	London	06	North London
000	London	07	GN/NE London Inner
000	London	61	Kent (LC & D)
000	London	62	Kent (SER)
000	London	63	Hastings Line
000	London	64	Oxted Line
000	London	65	Sussex Coast
000	London	66	Arun Valley
000	London	71	Portsmouth Lines
000	London	72	Southampton Lines
000	London	73	Windsor Lines
000	London	74	Wessex South
000	London	75	Salisbury-Exeter
000	London	81	Thames Valley
000	London	82	Chiltern
000	London	83	Northampton Lines
000	London	84	Thameslink North
000	London	85	GN Suburban
000	London	86	West Anglia
000	London	87	Gt. Eastern
000	London	88	LT & S

Table A.4
Origin-Destination Pairs: Non London Short Distance

Origin Zone		Destination Zone	
46	Norwich	86	West Anglia
30	West Mids inner	82	Chiltern
72	Southampton Lines	74	Wessex South
50	Cornwall	52	East Devon
15	Fylde	16	Lancaster
91	Far North	95	Edinburgh area
10	South Manchester	19	North Wales
94	W.Highland	96	Glasgow inner
36	Hereford/Worcester	81	Thames Valley
53	West Wales	55	Valleys
10	South Manchester	10	South Manchester
10	South Manchester	11	North Manchester
30	West Mids inner	30	West Mids inner
30	West Mids inner	32	West Mids outer
96	Glasgow inner	96	Glasgow inner
96	Glasgow inner	97	Glasgow outer

APPENDIX B. TESTING FOR COINTEGRATION

To test for cointegration, it is common practice first to establish the order of integration of the variables in question, and then - having established that the variables are of the same order of integration – to test whether there is at least one linear relationship among these variables.

To establish the order of integration of our explanatory variables, we have conducted panel unit root tests both for individual flows and for the whole panels. In general the individual flows test results suggest that the series are I(1) for all sectors and for our preferred set of explanatory variables. Furthermore, as the use of a panel increases the power of tests we apply the Im, Pesaran and Shin (1997) test for integration. [29] These results provide strong evidence that all our series are I(1).

Having established the order of integration of our series, we proceed with tests for co-integration. The literature observes a variety of possible ways in order to test for cointegration. Regarding the time series studies, it is found that extending the time series data length affects the order of integration and number of cointegrating vectors. However, testing for cointegration in panels is less pressing, because the spurious regression problem is reduced by the averaging. Some of the most popular cointegration tests for panel data are the McCoskey and Kao (1998) methodology and the Pendroni (1999) methodology. The assumption of heterogeneity between different flows remains while testing for cointegration in panel data.

McCoskey and Kao (1998) use a Lagrange Multiplier test on the residuals that takes the following form:

$$y_{i,t} = a_i + b_i x_{it} + e_{it} \qquad \text{(i)}$$

where,

$$e_{it} = q \sum_{j=1}^{t} u_{ij} + u_{it} \qquad \text{(ii)}$$

---

[29] The Im, Pesaran and Smith (1997) t-bar test averages the test statistics for the individual countries, and standardizes this average test statistic by its expected value and variance under the null hypothesis. The resulting standardized test statistic, denoted Ψ-bar statistic, is distributed as standard normal for large N, and its formula is:

$$\overline{\Psi} = \frac{\sqrt{N}(\bar{t}_{NT} - \frac{1}{N}\sum_{i=1}^{N} E[t_{iT}]}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} VAR[t_{iT}]}}$$

where t_{NT}-bar is the average of the N individual ADF test statistics and $E[t_{iT}]$ and $VAR[t_{iT}]$ are the empirical first and second moments of the ADF test statistics under the null.

McCoskey and Kao (1998) consider the above model under the null Ho: θ = 0, implying that there is cointegration in the panel, since for θ = 0, $e_{it} = u_{it}$ and the above regression is a system of cointegration. The alternative Hα: θ ≠ 0, is the lack of cointegration. The statistics are obtained by using the model that follows:

$$LM = \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{T=1}^{T} S_{it}^{+2}}{S^{+2}} \qquad \text{(iii)}$$

where S is the sum of the estimated error terms.

The estimation of the residuals can be applied by using OLS estimators and more specifically with the use of either FMOLS (Fully Modified OLS) or the DOLS (Dynamic OLS) estimator.

Pedroni (1999) also applies the cointegration test to the residuals of the regressions and uses the same Lagrange Multiplier test expressed in (iii) for a heterogeneous panel. Pedroni's (1999) approach refers in assuming trends for individual flows and in considering as null Ho: the lack of cointegration. In his analysis Pedroni (1999) proposes seven deferent cointegration statistics in order to capture the within and the between effects in his panel. All seven different statistics were calculated for our panels and the results indicated that in all cases, we have strong evidence of existence of cointegrating relationships.

APPENDIX C. DETAILED ESTIMATION RESULTS

C.1. Aggregate Results

Table C.1
Detailed Aggregate Estimation Results: London TCA to and from Rest of Country

	Period up to Hatfield		Entire period	
	Estimate	***t-statistic***	Estimate	***t-statistic***
Price	-0.64	***-22.9***	-0.64	***-26.1***
GDP	1.71	***67.5***	1.64	***82.3***
S_1	-0.09	***-8.7***	-0.09	***-8.6***
S_2	-0.03	***-2.5***	-0.03	***-3.0***
S_3	-0.05	***-4.2***	-0.05	***-4.5***
S_4	0.00	***-0.2***	0.00	***0.0***
S_5	-0.10	***-8.7***	-0.10	***-9.3***
S_6	-0.13	***-11.2***	-0.13	***-12.3***
S_7	-0.02	***-1.8***	-0.02	***-2.2***
S_8	0.06	***5.0***	0.06	***5.6***
S_9	0.08	***7.5***	0.08	***7.5***
S_10	-0.16	***-14.2***	-0.17	***-17.1***
S_11	-0.14	***-12.6***	-0.14	***-13.6***
S_12	-0.01	***-1.0***	0.00	***-0.4***
Hat_1	-	***-***	-0.25	***-8.1***
Hat_2	-	***-***	-0.16	***-6.0***

Source: NERA estimates

Table C.2

Detailed Aggregate Estimation Results: Non London Long Distance

	Period up to Hatfield		Entire period	
	Estimate	***t-statistic***	Estimate	***t-statistic***
Price	-0.82	***-37.8***	-0.82	***-41.8***
GDP	1.59	***61.5***	1.76	***85.1***
S_1	-0.12	***-11.1***	-0.12	***-11.8***
S_2	-0.09	***-8.8***	-0.09	***-9.6***
S_3	-0.10	***-9.5***	-0.10	***-10.1***
S_4	-0.11	***-10.7***	-0.11	***-10.7***
S_5	-0.16	***-14.8***	-0.15	***-14.8***
S_6	-0.15	***-14.4***	-0.15	***-14.8***
S_7	-0.05	***-4.7***	-0.04	***-4.1***
S_8	0.06	***5.7***	0.07	***7.0***
S_9	0.05	***4.6***	0.05	***5.1***
S_10	-0.15	***-14.2***	-0.17	***-17.6***
S_11	-0.13	***-12.6***	-0.14	***-14.2***
S_12	0.02	***1.8***	0.02	***2.2***
Hat_1	-	-	-0.28	***-16.5***
Hat_2	-	-	-0.15	***-9.0***

Source: NERA estimates

Table C.3
Detailed Aggregate Estimation Results: London TCA and South East

	Period up to Hatfield		Entire period	
	Estimate	***t-statistic***	Estimate	***t-statistic***
Price	-1.07	***-37.0***	-0.62	***-25.7***
GDP	1.71	***49.7***	1.21	***47.5***
EA	-0.04	***-0.5***	0.79	***10.3***
S_1	0.02	***3.5***	0.01	***1.6***
S_2	-0.02	***-3.9***	-0.01	***-1.3***
S_3	-0.02	***-3.8***	-0.02	***-3.5***
S_4	-0.02	***-3.6***	-0.02	***-2.6***
S_5	0.01	***1.8***	0.02	***3.1***
S_6	-0.04	***-6.8***	-0.03	***-4.4***
S_7	-0.03	***-4.8***	-0.01	***-2.2***
S_8	0.01	***2.2***	0.02	***3.6***
S_9	0.03	***5.8***	0.04	***6.5***
S_10	0.01	***1.7***	0.01	***1.2***
S_11	-0.09	***-14.4***	-0.09	***-15.1***
S_12	-0.02	***-3.7***	-0.04	***-6.6***
Hat_1	-	-	0.01	***1.3***
Hat_2	-	-	0.00	***0.2***

Source: NERA estimates

Table C.4
Detailed Aggregate Estimation Results: London TCA and South East (to London)

	Period up to Hatfield		Entire period	
	Estimate	***t-statistic***	Estimate	***t-statistic***
Price	-1.05	***-34.4***	-0.81	***-33.9***
GDP	1.54	***41.3***	1.28	***49.6***
EA	-0.03	***-0.4***	0.48	***6.4***
S_1	0.01	***1.9***	0.02	***2.9***
S_2	-0.01	***-1.5***	-0.01	***-1.4***
S_3	-0.07	***-11.7***	-0.07	***-12.6***
S_4	-0.03	***-5.2***	-0.04	***-6.3***
S_5	-0.03	***-4.4***	-0.02	***-2.8***
S_6	-0.02	***-3.1***	-0.02	***-2.9***
S_7	-0.05	***-8.0***	-0.03	***-5.9***
S_8	-0.04	***-6.0***	-0.04	***-5.9***
S_9	-0.02	***-3.8***	-0.02	***-4.2***
S_10	-0.02	***-2.5***	-0.01	***-1.3***
S_11	-0.04	***-6.2***	-0.04	***-7.1***
S_12	-0.07	***-10.1***	-0.05	***-9.0***
Hat_1	-	-	0.05	***5.4***
Hat_2	-	-	0.06	***7.2***

Source: NERA estimates

Table C.5
Detailed Aggregate Estimation Results: London TCA and South East (from London)

	Period up to Hatfield		Entire period	
	Estimate	***t-statistic***	Estimate	***t-statistic***
Price	-0.99	***-18.4***	-0.51	***-11.8***
GDP	2.19	***48.2***	1.78	***51.4***
EA	-1.09	***-8.5***	-0.52	***-4.1***
S_1	0.02	***1.7***	0.01	***1.0***
S_2	0.03	***2.8***	0.02	***1.6***
S_3	-0.05	***-4.1***	-0.08	***-7.3***
S_4	-0.01	***-0.4***	-0.03	***-2.7***
S_5	0.05	***4.3***	0.04	***3.4***
S_6	0.07	***5.9***	0.05	***4.8***
S_7	0.03	***2.3***	0.02	***1.5***
S_8	0.04	***3.6***	0.02	***2.2***
S_9	0.04	***3.3***	0.04	***3.8***
S_10	0.05	***4.1***	0.05	***5.0***
S_11	0.06	***4.8***	0.06	***5.1***
S_12	-0.02	***-1.3***	-0.01	***-1.1***
Hat_1	-	-	0.07	***4.0***
Hat_2	-	-	0.09	***5.6***

Source: NERA estimates

Table C.6
Detailed Aggregate Estimation Results: Non London Short Distance

	Period up to Hatfield		Entire period	
	Estimate	***t-statistic***	Estimate	***t-statistic***
Price	-0.99	***-20.7***	-0.95	***-22.5***
GDP	1.35	***31.9***	1.34	***40.6***
S_1	0.02	***1.1***	0.01	***0.6***
S_2	-0.01	***-0.3***	-0.01	***-0.4***
S_3	0.01	***0.9***	0.01	***0.4***
S_4	0.06	***3.2***	0.05	***3.0***
S_5	0.08	***4.5***	0.07	***4.3***
S_6	0.12	***7.1***	0.11	***7.2***
S_7	0.13	***7.3***	0.12	***7.6***
S_8	0.09	***5.2***	0.08	***5.5***
S_9	0.09	***5.3***	0.09	***5.6***
S_10	0.07	***3.9***	0.06	***4.2***
S_11	-0.07	***-4.2***	-0.08	***-5.2***
S_12	-0.08	***-4.9***	-0.09	***-6.2***
Hat_1	-	-	-0.11	***-4.9***
Hat_2	-	-	-0.05	***-2.1***

Source: NERA estimates

C.2. Disaggregate Results

Note: all disaggregate results refer to the entire period, i.e. including the Hatfield accident and subsequent periods.

Table C.7
Detailed Disaggregate Estimation Results: London TCA to and from Rest of Country

	Fully flexible		Reduced/restricted	
	Estimate	t-statistic	Estimate	t-statistic
Price FS	-0.49	***-12.0***	0.26	***8.4***
price RO	0.15	***3.7***	-0.82	***-25.2***
GDP	0.84	***19.6***	1.78	***55.8***
S_1	0.09	***5.7***	-0.02	***-1.8***
S_2	0.14	***8.8***	-0.03	***-2.5***
S_3	-0.08	***-5.0***	-0.01	***-0.6***
S_4	0.00	***-0.2***	0.00	***0.1***
S_5	0.06	***3.7***	-0.02	***-1.5***
S_6	0.07	***4.4***	0.02	***1.6***
S_7	0.02	***1.4***	0.04	***3.3***
S_8	0.03	***1.7***	0.07	***5.7***
S_9	0.09	***6.0***	-0.03	***-3.0***
S_10	0.06	***3.8***	-0.04	***-3.5***
S_11	0.07	***4.2***	0.00	***0.0***
S_12	-0.06	***-3.9***	0.00	***-0.2***
Hat_1	-0.19	***-3.7***	-0.27	***-7.5***
Hat_2	-0.09	***-1.9***	-0.23	***-7.0***

Source: NERA estimates

Table C.8
Detailed Disaggregate Estimation Results: Non London Long Distance

	Fully flexible		Reduced/restricted	
	Estimate	t-statistic	Estimate	t-statistic
Price FS	-1.16	***-28.8***	0.48	***17.7***
price RO	0.74	***13.3***	-0.70	***-19.6***
GDP	1.90	***37.6***	1.29	***37.1***
S_1	-0.09	***-4.4***	-0.07	***-5.1***
S_2	-0.10	***-4.5***	-0.07	***-5.3***
S_3	-0.03	***-1.6***	-0.03	***-2.1***
S_4	-0.04	***-2.0***	-0.05	***-3.8***
S_5	-0.10	***-4.5***	-0.08	***-5.5***
S_6	-0.03	***-1.5***	-0.02	***-1.7***
S_7	0.04	***1.9***	0.05	***3.4***
S_8	0.04	***1.8***	0.05	***3.7***
S_9	-0.02	***-1.2***	-0.14	***-9.8***
S_10	-0.16	***-7.6***	-0.10	***-7.2***
S_11	-0.04	***-1.8***	0.00	***0.0***
S_12	0.01	***0.5***	0.07	***4.8***
Hat_1	0.09	***2.7***	0.07	***3.0***
Hat_2	0.16	***4.1***	0.14	***5.3***

Source: NERA estimates

Table C.9
Detailed Disaggregate Estimation Results: London TCA and South East

	Seasons		Non-seasons	
	Estimate	t-statistic	Estimate	t-statistic
Price S	-0.83	***-29.6***	-0.24	***-5.3***
Price N_S	0.50	***11.6***	-0.40	***-16.0***
GDP	0.50	***12.6***	1.90	***66.4***
EA	0.95	***9.4***	-	-
S_1	0.01	***1.2***	-0.04	***-4.9***
S_2	-0.08	***-10.8***	-0.04	***-5.3***
S_3	-0.09	***-12.4***	-0.03	***-4.1***
S_4	-0.03	***-4.2***	0.03	***3.5***
S_5	0.01	***0.7***	-0.01	***-1.5***
S_6	-0.03	***-4.1***	-0.04	***-4.9***
S_7	-0.05	***-7.0***	-0.02	***-3.3***
S_8	-0.06	***-8.0***	0.03	***4.0***
S_9	-0.05	***-6.8***	0.03	***4.4***
S_10	-0.03	***-3.9***	-0.14	***-17.1***
S_11	-0.09	***-12.7***	-0.09	***-12.0***
S_12	-0.07	***-9.3***	-0.04	***-5.6***
Hat_1	0.04	***3.9***	-0.07	***-5.9***
Hat_2	0.06	***5.3***	-0.01	***-0.8***

Source: NERA estimates

Table C.10
Detailed Disaggregate Estimation Results: Non London Short Distance

	Seasons		Non-seasons	
	Estimate	t-statistic	Estimate	t-statistic
Price S	-1.50	***-20.7***	0.12	***4.3***
Price N_S	-0.42	***-4.4***	-1.21	***-27.0***
GDP	2.52	***33.0***	1.23	***32.1***
S_1	-0.05	***-1.5***	0.00	***-0.1***
S_2	-0.15	***-4.7***	0.01	***0.4***
S_3	-0.29	***-9.2***	-0.09	***-5.6***
S_4	-0.10	***-3.1***	-0.09	***-5.1***
S_5	-0.10	***-3.0***	-0.02	***-1.4***
S_6	-0.19	***-5.8***	-0.01	***-0.5***
S_7	-0.23	***-7.0***	-0.03	***-1.6***
S_8	-0.24	***-7.4***	-0.01	***-0.8***
S_9	-0.14	***-4.4***	0.00	***-0.1***
S_10	-0.19	***-5.9***	0.01	***0.5***
S_11	-0.25	***-7.6***	0.04	***2.3***
S_12	-0.20	***-6.3***	0.02	***1.2***
Hat_1	0.02	***0.4***	0.03	***1.3***
Hat_2	0.08	***1.7***	0.07	***3.1***

Source: NERA estimates

APPENDIX D. MULTICOLLINEARITY

The results obtained in the estimation of the disaggregated elasticities and cross-elasticities must be interpreted with care. This is because of the presence of multicollinearity among the measured explanatory variables that have been used for the disaggregate analysis of the four panels. Multicollinearity indicates that the explanatory variables are too highly intercorrelated (i.e. there exists a strong linear relationship among some or all of the explanatory variables). The presence of multicollinearity generates a lack of accuracy in estimating the effects that the explanatory variables produce on the dependent variable (i.e. elasticities of the model). If this data problem occurs, the regression analysis is not able to precisely allocate the impact of the explanatory variables on the dependent variable, since some or all of those explanatory variables are approximately similar to a linear combination of the other explanatory variables. In other words, if multicollinearity exists, the estimates of the coefficients may have the unexpected sign or an implausible magnitude.

This data problem cannot be overcome easily. An approach that is sometimes pursued is dropping one or more of the regressors that produce the multicollinearity. We have not been able to do this in the context of our work, firstly because one of the objectives of the project was to try to estimate the cross elasticities, and secondly because we will demonstrate that the potential multicollinearity problem may be due to the correlation between the series of the ticket prices.

The rest of the Appendix will provide evidence on the existence of multicollinearity in the four panels. In order to do so, three measures have been calculated for the four panels. These are:

- The correlation coefficients (ρ)[30] between the different pairs of regressors. This coefficient takes values in the numerical range [-1,1]. The closer to 1 the absolute value of the ρ is, the stronger the evidence of multicollinearity is;

[30] The ρ is the ratio between the (i) covariance between the pair of regressors, and (ii) product of their standard deviations. The coefficient measures how strong is the linear relationship between the pair of regressors.

- The Variation Inflation Factor (VIF)[31] for the different regressors. This factor takes values in the numerical range [1,+∞]. The higher the value of the VIF is, the stronger the evidence of multicollinearity is;[32] and
- The measure of tolerance (TOL)[33] for the different regressors. This measure takes values in the numerical range [0,1]. The closer the value of the TOL to 0 is, the stronger the evidence of multicollinearity is.

Table.D.1 to Table D.4 below present these measures for the relevant regressors in the four panels. The tables show that the main source of the multicollinearity problem is related to the linear relationship between the series of alternative ticket types. For these variables in the four panels, the ρ's take a value higher than 0.80; the VIF's are well above 1; and the TOL's are close to zero. The values of the VIF for the two ticket type variables in the London TCA and South East panel and the Non London Long Distance panel are particularly remarkable, with values well above 10.

---

31 The VIF is a measure that can guide in identifying multicollinearity. To develop this concept it is useful to note that the variance of the OLS estimator for a typical regression coefficient (say β_i) can be shown to be the following:

$$Var(\beta_i) = \sigma^2 / S_{ii} (1-R_i^2)$$

Where (i) σ^2 is the variance of the error term of the model, (ii) S_{ii} is the sum of the differences to the square between the observations of the explanatory variable i and its arithmetic average, and (iii) R_i^2 is the unadjusted R^2 when the explanatory variable i is regressed against the other explanatory variables in the model.

The intuitive insight of this expression is that the stronger the linear relationship between the explanatory variables (i.e. multicollinearity) is, the larger the variance of the estimates of the regressions coefficients (and consequently the more difficult to obtain an accurate estimate of the true value of the parameter). In particular, if R_i^2 is equal to 0, then the variance of the estimation takes its minimum possible value, which is σ^2 / S_{ii}. Dividing this last expression into the above expression for $Var(\beta_i)$, the VIF is obtained.

$$VIF = 1 / (1-R_i^2)$$

It is straightforward to understand that the higher the VIF is, the lower the accuracy of the estimates of the parameter.

32 There is not a defined threshold for this factor as it is not based on any probability distribution. Nevertheless, as a rule of thumb, if the VIF exceeds 10 (i.e. the R^2 of the ***auxiliary regressions*** exceeds 0.90), it is said that there is strong evidence of multicollinearity.

33 The TOL is defined as:

$$TOL(\beta_i) = 1 / VIF = 1- R_i^2$$

It is clear to observe that the lower the TOL is, the lower the accuracy of the estimates of the parameter.

Table.D.1
Evidence of Multicollinearity: London TCA to and from Rest of Country

ρ	Fully Flexible Tickets Prices and Reduced+Others Ticket Prices	0.83
	Fully Flexible Tickets Prices and GDP	0.20
	Reduced+Others Ticket Prices and GDP	0.02
VIF	Fully Flexible Tickets Prices	3.68
	Reduced+Others Ticket Prices	3.53
	GDP	1.13
TOL	Fully Flexible Tickets Prices	0.27
	Reduced+Others Ticket Prices and	0.28
	GDP	0.89

Table.D.2
Evidence of Multicollinearity: Non London Long Distance

ρ	Fully Flexible Tickets Prices and Reduced+Others Ticket Prices	0.97
	Fully Flexible Tickets Prices and GDP	0.02
	Reduced+Others Ticket Prices and GDP	0.10
VIF	Fully Flexible Tickets Prices	17.29
	Reduced+Others Ticket Prices	17.27
	GDP	1.01
TOL	Fully Flexible Tickets Prices	0.06
	Reduced+Others Ticket Prices and	0.06
	GDP	0.99

Table.D.3
Evidence of Multicollinearity: London TCA and South East

ρ	Non-Seasonal Tickets Prices and Seasonal Ticket Prices	0.97
	Seasonal Tickets Prices and GDP	0.12
	Non-Seasonal Ticket Prices and GDP	0.10
VIF	Seasonal Tickets Prices	15.40
	Non-Seasonal Ticket Prices	15.35
	GDP	1.02
TOL	Seasonal Tickets Prices	0.07
	Non-Seasonal Ticket Prices	0.07
	GDP	0.98

Table D.4
Evidence of Multicollinearity: Non London Short Distance

ρ	Non-Seasonal Tickets Prices and Seasonal Ticket Prices	0.92
	Seasonal Tickets Prices and GDP	0.02
	Non-Seasonal Ticket Prices and GDP	0.07
VIF	Seasonal Tickets Prices	6.65
	Non-Seasonal Ticket Prices	6.65
	GDP	1.01
TOL	Seasonal Tickets Prices	0.15
	Non-Seasonal Ticket Prices	0.15
	GDP	1.00

APPENDIX E. PREDICTION TESTS IN EVIEWS

EViews, the software package that we have used for the present study, is a typical professional econometrics package. In the present Appendix, we provide a simple example to show how it could be used as part of the updating mechanism (see Chapter 7)

For the purposes of our simple example, we have created an equation and estimated a (not very well justified) model with GDP as a function of the money supply. The results are shown in Table E.1 below.

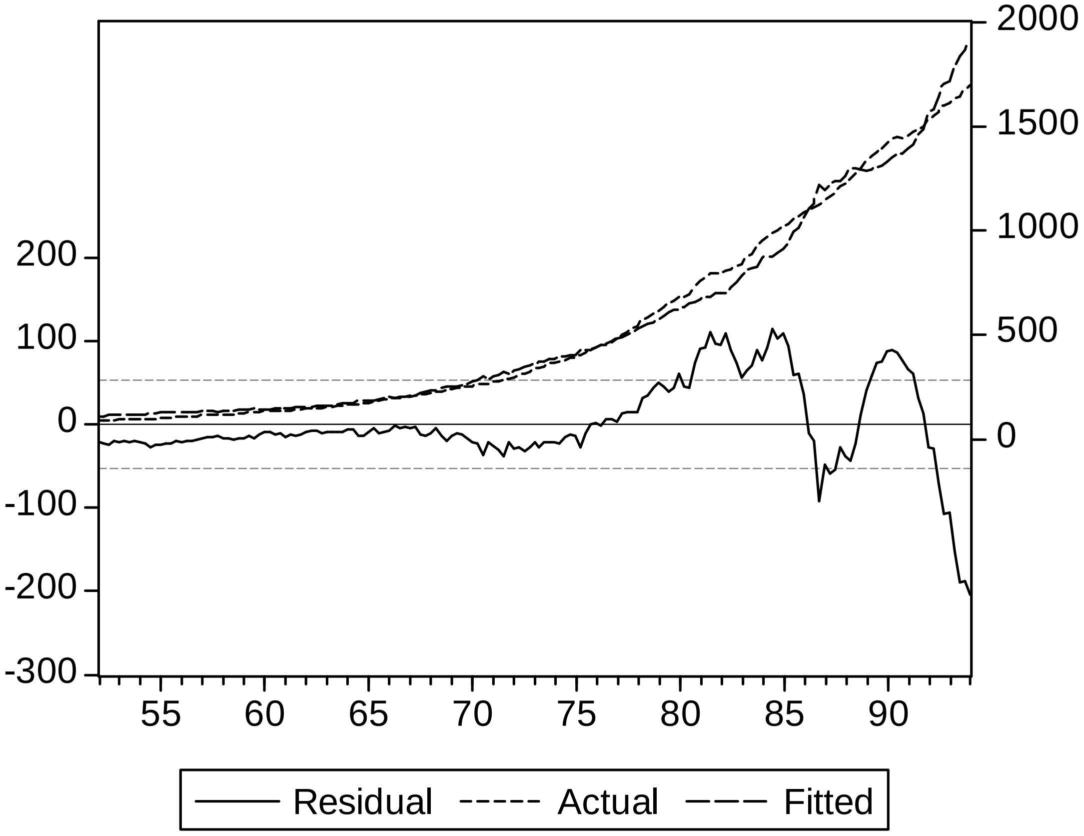
Table E.1
Results of Illustrative Example

Dependent Variable: GDP
Method: Least Squares
Date: 11/26/03 Time: 09:46
Sample: 1952:1 1994:1
Included observations: 169

Variable	Coefficient	Std. Error	t-Statistic
C	-99.86719	6.777091	-14.73600
M1	1.654676	0.013778	120.0987
R-squared	0.988554	Mean dependent var	
Adjusted R-squared	0.988486	S.D. dependent var	
S.E. of regression	52.41340	Akaike info criterion	
Sum squared resid	458776.4	Schwarz criterion	
Log likelihood	-907.8931	F-statistic	
Durbin-Watson stat	0.072879	Prob(F-statistic)	

We can graph the equation using View/Actual, Fitted, Residual and then choosing Actual, Fitted, Residual Graph. The resulting graph is reproduced here as Figure E.1.

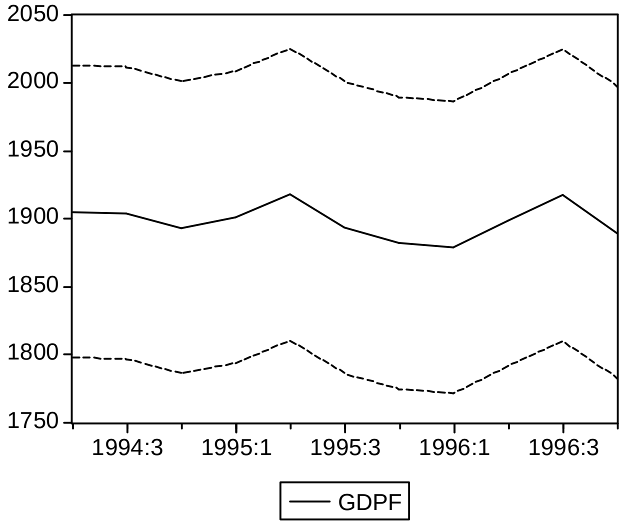
Figure E.1
Graph of Illustrative Example



This shows that the equation fits badly towards the end of the period.

We then push the forecast button on the Equation toolbar. This provides us with a graph of the forecasts and the plus and minus two standard error bands, shown in Figure E.2. When the outturn values lie outside the predefined error bands, an error message could be triggered suggesting that the model elasticities should be reviewed.

Figure E.2
Forecasting Graph for Illustrative Example



Forecast: GDPF
Actual: GDP
Forecast sample: 1994:2 1996:4
Included observations: 11

Root Mean Squared Error	97.67580
Mean Absolute Error	77.94762
Mean Abs. Percent Error	4.366969
Theil Inequality Coefficient	0.026177
Bias Proportion	0.461566
Variance Proportion	0.331591
Covariance Proportion	0.206843

APPENDIX F. REFERENCES

Bentzen, J (1994) "An empirical analysis of gasoline demand in Denmark using cointegration techniques" ***Energy Economics,*** Vol 16, No 2, p139-143.

Eltony, M N, Al-Mutairi, N H (1995) "Demand for gasoline in Kuwait: An empirical analysis using cointegration techniques" ***Energy Economics,*** Vol 17, No 3, p249-253.

Espey, M (1998) "Gasoline demand revisited: An international meta-analysis of elasticities" ***Energy Economics***, Vol 20, p 273-295.

Dargay, J M, Hanly, M (2002) "The demand for local bus services in England" ***Journal of Transport Economics and Policy,*** Vol 36, No 1 (January), p 73-91.

Graham, D J, Glaister, S, (2002) "The demand for automobile fuel" ***Journal of Transport Economics and Policy,*** Vol 36, No 1 (January), p 73-91.

NERA (1999) ***Analysis of Passenger Rail Demand: A Report for OPRAF***

Owen, A D, Phillips, G D A (1987) "The characteristics of railway passenger demand" ***Journal of Transport Economics and Policy,*** Vol 21 (September), p231-253.

Ramanathan, R (1999) "Short- and long-run elasticities of gasoline demand in India: An empirical analysis using cointegration techniques" ***Energy Economics***, Vol 21, p321-330.

Samimi, R (1995) "Road transport energy demand in Australia: A cointegration approach" ***Energy Economics,*** Vol 17, No 4, p329-339.