

Department for Transport

**Research Into The Appraisal Of
Long Term Benefits Of Transport
Schemes**

Final Report

Issue 1

F1 | 7 November 2016

This report takes into account the particular instructions and requirements of our client.

It is not intended for and should not be relied upon by any third party and no responsibility is undertaken to any third party.

Job number 248185-00

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



ARUP

Document Verification



Job title		Research Into The Appraisal Of Long Term Benefits Of Transport Schemes		Job number		248185-00	
Document title		Final Report		File reference			
Document ref		Issue 1					
Revision	Date	Filename	Long Term Benefits Report - Draft 3 v48.docx				
Draft Final	15 Aug 2016	Description	Formal draft for DfT				
			Prepared by	Checked by	Approved by		
		Name	Doug Rose, Kevin Yeung (both Arup), John Bates, Tom Worsley, Richard Batley (all ITS Leeds)	Margot Finley (Arup), Peter Mackie (ITS Leeds)	Matt Dillon (Arup)		
		Signature					
Draft Final 2	19 Sep 2016	Filename	Long Term Benefits Report - Draft v56.docx				
		Description	Formal draft 2 for DfT				
			Prepared by	Checked by	Approved by		
		Name	Doug Rose, Kevin Yeung (both Arup), John Bates, Tom Worsley, Richard Batley (all ITS Leeds)	Margot Finley (Arup), Peter Mackie (ITS Leeds)	Matt Dillon (Arup)		
	Signature						
Draft Final 3	12 Oct 2016	Filename	Long Term Benefits Report - Draft v60.docx				
		Description	Formal draft 3 for DfT				
			Prepared by	Checked by	Approved by		
		Name	Doug Rose, Kevin Yeung (both Arup), John Bates, Tom Worsley, Richard Batley (all ITS Leeds)	Margot Finley (Arup), Peter Mackie (ITS Leeds)	Matt Dillon (Arup)		
	Signature						

Final Issue 1	7 Nov 2016	Filename	Long Term Benefits Report – Final – ISSUE1.docx			
		Description	Final Report			
			Prepared by	Checked by	Approved by	
		Name	Doug Rose, Kevin Yeung (both Arup), John Bates, Tom Worsley, Richard Batley (all ITS Leeds)	Margot Finley (Arup), Peter Mackie (ITS Leeds)	Matt Dillon (Arup)	
		Signature				

Issue Document Verification with Document



Contents

	Page
Executive Summary	7
1 Introduction	11
1.1 Overview	11
1.2 What Drives Long Term Benefits?	12
1.3 Demand Cap	14
1.4 Limitations Of The Demand Cap	15
1.5 Uncertainty	18
1.6 Theoretical Background	19
1.7 Certainty Equivalence	24
1.8 Study Approach	25
2 Historic Demand Trends	26
2.1 Transport	26
2.2 Rail	29
2.3 Rail Market Saturation	36
3 Historic Scheme Trends	39
3.1 Scheme Selection	39
3.2 Scheme Demand Profiles	40
3.3 Scheme Demand Growth vs Population Growth	42
4 Simulation Model	43
4.1 Simulation Model	43
4.2 Project Case Study	70
4.3 Conclusions	71
5 Long Term Demand Modelling	73
5.1 Strategic Models And Cross Sectional Data	73
5.2 Practical Use of Strategic Models	74
5.3 Long Term Use of Elasticities	75
6 Discussion	80
6.1 Is the demand cap a valid concept in a literal sense?	80
6.2 Is a demand cap a fair simple proxy for a complex reality?	81
6.3 How should we deal with uncertainty?	82
6.4 Our Proposal	83
6.5 Peer Review	85
7 Recommendations	86
7.1 20 Years Modelling from Opening	86

7.2	Adopt a ‘Certainty Equivalence’	87
7.3	Remove the Demand Cap	88

Tables

Table 1 – Green Book Discount Rates
Table 2 – Schemes examined in summary analysis
Table 3 – Schemes with significant further investment
Table 4 – Mean annual demand growth rates by age of project and mode
Table 5 – Tests considered
Table 6 – 2070 DM and DS Equilibrium Demand and Generalised Time
Table 7 – Values for Example Rule of Half Calculation
Table 8 – Highway PVB for Demand Cap with Years-based Approach
Table 9 – Highway PVB for Income Only Years-based Demand Cap
Table 10 – Highway PVB for a Generalised Cost-based Demand Cap
Table 11 – Highway PVB for Uncertainty Equivalent Test
Table 12 – 2070 DM and DS Equilibrium Demand and Generalised Journey Time
Table 13 – Rail PVB for Demand Cap using Years-based Approach
Table 14 – Rail PVB for GDP / Employment only Years-based Demand Cap
Table 15 – Rail PVB for Generalised Cost-based Demand Cap
Table 16 – Rail PVB for Uncertainty Equivalent Test
Table 17 – Full list of schemes examined as part of this project
Table 18 – Growth Factor by Vehicle Class to obtain 2060 Travel Demand Matrices
Table 19 – Model convergence statistics
Table 20 – Net Present Value and Benefit Cost Ratio with and without inclusion of Modelled Year 2060

Figures

Figure 1 – Simplified benefit-cost ratio drivers
Figure 2 – England Traffic forecasts
Figure 3 – Demand profile for a hypothetical rail scheme
Figure 4 – Discounted benefits profile for a hypothetical* scheme
Figure 5 – Illustration of Supply-Demand equilibrium over time
Figure 6 – Illustration of Benefits over time
Figure 7 – Passenger kilometres by mode in Great Britain, 1952-2014
Figure 8 – Trends in trips, distance and time spent travelling by a typical person – indexed

- Figure 9 – DfT central or mid-point car traffic projections and outturn statistics**
- Figure 10 – Estimate of rail trips per person p.a. in Great Britain, 1831 to 2014**
- Figure 11 – Rail journeys and length of National Rail route in Great Britain, 1831 to 2014**
- Figure 12 – Rail trips and car ownership**
- Figure 13 – Approximate change in ticket purchasing locations over time for trips on long distance rail operators, 2003-2015**
- Figure 14 – Percentage of households in England with access to a car, 1951 to 2014**
- Figure 15 – Relationship between rail passenger journeys and GDP**
- Figure 16 – Median growth profile for all schemes sampled**
- Figure 17 – 2010 Base Year Demand**
- Figure 18 – 2010, 2025 and 2040 Highway Demand Curves**
- Figure 19 – Highway Demand for a Fixed Cost (1.0 min/km)**
- Figure 20 – 2010 Highway Supply Curve**
- Figure 21 – Highway Supply Curves for 2010 and 2040**
- Figure 22 – 2010 Highway DM and DS Supply Curves**
- Figure 23 – 2010 Highway Demand, DM Supply and DS Supply Curves**
- Figure 24 – 2070 Highway Demand, DM Supply and DS Supply Curves**
- Figure 25 – Highway Do Minimum and Do Something Equilibrium Paths**
- Figure 26 – Example Highway Discounted Benefits over the Appraisal Period**
- Figure 27 – Equilibrium Paths for Uncapped Highway Demand**
- Figure 28 – Discounted Benefits for Uncapped Highway Demand**
- Figure 29 – Highway Demand and Supply Curves for a 20 year cap**
- Figure 30 – Highway Discounted Benefits for Years-based Demand Cap**
- Figure 31 – Highway Discounted Benefits for Income Only Years-based Demand Cap**
- Figure 32 – Highway Demand and Supply Curves for a 150% GC Demand Cap**
- Figure 33 – Highway Discounted Benefits for Generalised Cost-based Demand Cap**
- Figure 34 – Highway Discounted Benefits for Certainty Equivalent Test**
- Figure 35 – Example Rail Demand Curves for 2010, 2025 and 2040**
- Figure 36 – Rail DM (8-car) and DS (12-car) Supply Curves**
- Figure 37 – Change in Rail Demand - Supply Curves and Equilibrium**
- Figure 38 – Rail Discounted Benefits over the Appraisal Period (Uncapped Demand)**
- Figure 39 – Rail Demand and Supply Curves for 20-Year Demand Cap**
- Figure 40 – Rail Discounted Benefits for Years-based Demand Cap**
- Figure 41 – Rail Discounted Benefits for GDP / Employment Years-based Demand Cap**
- Figure 42 – Change in Rail Demand-Supply Curves with a Generalised Cost Cap**

- Figure 43 – Rail Discounted Benefits for Generalised Cost-based Demand Cap**
- Figure 44 – Rail Discounted Benefits for Certainty Equivalent Test**
- Figure 45 – Growth in passengers for individual rail transport schemes indexed to the 20th year after opening**
- Figure 46 – Growth in passengers for individual road transport schemes indexed to the 20th year after opening**
- Figure 47 – Median growth profile for all schemes sampled**
- Figure 48 – Indexed growth in patronage 20 years after transport schemes was completed**
- Figure 49 – Passenger journeys on Eurostar 1995 to 2014**
- Figure 50 – Passenger journeys and train kilometres on the InterCity East Coast route, 1997 to 2015.**
- Figure 51 – Passenger journeys and vehicle miles on Glasgow Subway 1980 to 2015**
- Figure 52 – Glasgow population 1981-2014**
- Figure 53 – Growth in patronage 20 years after opening compared to growth in local GVA/capita and population - Re-opening of Glasgow Subway**
- Figure 54 – Passenger journeys and vehicle kilometres on DLR 1988 to 2015**
- Figure 55 – Passenger journeys per vehicle kilometre on the DLR, 1988 to 2015**
- Figure 56 – Passenger journeys on Tyne and Wear Metro 1980 to 2015**
- Figure 57 – Tyne and Wear (Met County) population 1981 to 2014**
- Figure 58 – Growth in patronage 20 years after opening compared to growth in local GVA/capita and population - Tyne and Wear Metro**
- Figure 59 – Passenger journeys and vehicle kilometres on Manchester Metrolink 1992 to 2015**
- Figure 60 – Approximate number of passengers boarding onto Victoria Line services on a typical weekday and operated KMs on the line, 1998 to 2014**
- Figure 61 – Approximate number of passengers boarding onto Jubilee Line services on a typical weekday and operated KMs on the line, 1998 to 2014**
- Figure 62 – Approximate number of passengers boarding rail at Stratford 1981 to 2015**
- Figure 63 – Approximate number of passengers boarding rail at Canary Wharf, 2000-2015**
- Figure 64 – Morning peak travel to the Isle of Dogs (including Canary Wharf) by mode of transport, 1988 to 2013**
- Figure 65 – Employment in Tower Hamlets 2004 to 2015**
- Figure 66 – Passenger journeys and vehicle kilometres on Blackpool Tram, 1963 to 2-15**
- Figure 67 – Passenger journeys and vehicle kilometres on Blackpool Tram, 1963 to 2015**
- Figure 68 – Blackpool population 1981 to 2015 (Note: data before 1983 was not available)**
- Figure 69 – Total overnight stays in Blackpool from people on holiday or visiting friends/relatives, 2006 to 2015**

- Figure 70 – Vehicles crossing at Dartford Crossing (tunnel or bridge) per annum, 1961 to 2015**
- Figure 71 – Percentage of households in England with access to a car 1961 to 2015**
- Figure 72 – Approximate growth in traffic on the M3 and all roads, 1995 to 2015**
- Figure 73 – Approximate growth in traffic on the M11 and all roads, 1993 to 2015**
- Figure 74 – Approximate growth in traffic on M25 and all roads, 1993 to 2015**
- Figure 75 – Approximate growth in traffic on M11, 1999 to 2015**
- Figure 76 – Growth in at the A19 Barlby Junction bypass (near York)**
- Figure 77 – Growth in traffic at the A19 Billingham Diversion (near Middlesbrough)**
- Figure 78 – Growth in traffic after the dualling of the A66 between Banks Gate and Bowes West**
- Figure 79 – Growth in traffic at the Stocksbridge bypass on the A616 (near Sheffield)**
- Figure 80 – Growth in patronage 20 years after opening compared to growth in local GVA/capita and population - Stocksbridge bypass on the A616 (near Sheffield)**
- Figure 81 – Growth in traffic on the M74 between J12 and J13**
- Figure 82 – Growth in traffic on the M65 between Nelson to Colne (J13 to J14)**
- Figure 83 – Growth in traffic at on the M42 Southern Links (M5 to M42 J1)**
- Figure 84 – Growth in traffic on the M23 between Hooley and Mertsham (J7-J8)**
- Figure 85 – Growth in traffic on the M25 between Micklefield to South Mimms (J19 to J23)**
- Figure 86 – Growth in traffic on the M3 between Sunbury to Lightwater (J1 to J3)**
- Figure 87 – Growth in traffic on the M5 around Avonmouth Bridge (J18 to J19)**
- Figure 88 – Growth in traffic on the M40 between Stokenchurch to Waterstock (J5 to J8a)**
- Figure 89 – Indices of Trips, Total Travel Distance, Total Travel Time and Average Network Speed by Modelled Year**
- Figure 90 – Profile of Discounted Benefits over Appraisal Period (£m 2010 prices)**
- Figure 91 – A Typical Choice Hierarchy and Associated Cost Transfers**

Appendices

Appendix A

Demand Case Studies

Appendix B

Road Case Study

Executive Summary

Arup and the Institute for Transport Studies (ITS), University of Leeds were commissioned by the Department for Transport (DfT) to undertake research into the appraisal of long term economic benefits of transport schemes.

Study Purpose

The intention of this study was to advance the discussion around the quantification of long term benefits in relation to the demand cap. We were asked to consider a number of questions concerning long term demand forecasting, accounting for uncertainty, the profile of benefits throughout the appraisal period and other elements of benefits appraisal. We were asked to provide recommendations where there is evidence to do so. The scope did not include producing a definitive solution or developing guidance for WebTAG. This study considers the core transport user benefits considered in standard appraisal and does not consider Wider Economic Benefits.

Drivers of Benefits

The main drivers of benefits are the growth in **demand** as incomes and population increase, which also leads to growth in the value of the time savings (and congestion relief). Much of this report focuses on long term travel demand patterns, particularly beyond 20 years of scheme opening, and the use of the demand cap. The focus is on the use of the demand cap rather than the value of time itself, as this is already a well-researched area.

In general, as demand increases, more congestion and crowding is likely to occur. Thus in considering how benefits might evolve in the longer term, the interaction between the supply of capacity and demand has been a key consideration.

Long Term Travel Trends

Considering the importance of demand projections to the quantification of benefits, we have reviewed historic demand trends in transport in general, and in particular for rail, and have also considered the demand profile of a number of individual schemes.

We have considered whether the evidence on travel demand shows any support for the view that growth might be reducing to a rate at which long term benefits might cease to increase. Our study suggests that rail demand shows no clear signs of a reduction in rates of growth, while trends in car travel over the past decade can be explained by changes in fuel costs and in household incomes rather than a cessation in overall demand growth.

Additionally, we considered the use of elasticities for rail far into the future and found no conclusive evidence that they should be used for any specific period of time. However, the GDP elasticity remained consistent over many years indicating that its long term use may be valid.

Individual Scheme Trends

We also analysed a number of transport schemes for which data on traffic and passenger flows over a long period are available. Many of the schemes had demand growth after the 20th year from opening (and hence beyond the 20th year from the original appraisal), although some of these schemes that experienced the highest levels of growth had benefitted from additional investment in capacity which might have enabled growth to continue.

Discount Rate and Introduction of the Cap

Prior to 2002 the appraisal of transport schemes was carried out over a 30 year period using discount rates which changed from time to time, but which generally exceeded 6%. Forecasts were made over a 30 year period, after appropriate allowance for uncertainty, and the discount rate reduced the weight put on more distant and uncertain benefits.

The reduction in the discount rate (3.5% for the first 30 years, followed by 3% over the next 30 years) and an increase in the appraisal period to 60 years in 2002 increased the importance of long term benefits to the overall appraisal. The DfT adopted a pragmatic approach with the aim of simplifying the appraisal process and reflecting uncertainty in long term forecasting, and imposed a demand cap. Beyond the year in which the cap was applied – usually 20 years from the date of the appraisal being undertaken – demand was assumed to remain constant and benefits grew in line with the assumed increase in the value of time savings.

Strengths of the Demand Cap

The primary strength of the demand cap was that it accounted for long term uncertainty in demand and benefits. Other strengths included that it was easy to implement, easy to understand and it meant that practitioners did not have to worry about the complexities of long term demand forecasting and use of transport models.

Limitations of Demand Cap

A limitation of this pragmatic approach became clear when schemes that differed in the quantity of additional capacity provided were compared. The demand cap effectively put no value on capacity which was not fully utilised in the cap year even if larger increments to capacity were a more effective way of providing for predicted longer term growth.

Consideration of Alternatives to a Demand Cap

It could be said that the ideal solution is to forecast long term demand by running the transport model used to estimate the benefits prior to the cap, over the whole of the life of the scheme. However, this presents a number of issues as outlined below.

- In many cases, it is not **practical** to recommend running transport models far into the future due to cost, model run times, challenges with defining do-minimum and model convergence. The transport model can become unusable in the context of continuing demand growth. As ever increasing

demand is fed on to the networks, the processes in the model which bring demand and supply into equilibrium have an increased likelihood of failing to operate effectively.

- **Uncertainty** about the drivers of demand and the relationship between these drivers and transport user benefits reduces the reliance that might be placed on forecasts so far into the future. The likelihood of changes in technology and in policy increases this uncertainty.

Recommendations

Based on the work of this study, we recommend that DfT considers and undertakes further research on the following three items. These are suggested as a potential alternative to the demand cap, but it is not recommended that these are implemented without further work.

1. Specifying that the 20 year complex modelling period should run from the **predicted year of scheme opening** rather than 20 years from the date of the appraisal. This may be achievable from a transport modelling perspective, reflects the concept of a “design year” 20 years from opening and provides a higher degree of fairness between schemes.
2. Taking specific account of uncertainty through use of a **certainty equivalence**/cost of risk which would apply to the entire appraisal period and could be implemented simply by a table in WebTAG.
3. Developing **guidance** for practitioners around the method for quantifying long term benefits beyond the initial 20-year period as a result of the removal of the demand cap.

We see these three recommendations as being considered together rather than as options to be considered individually.

On point (3), we propose two options:

- **Option A** – Extrapolate demand from the 20-year complex modelling period based on the model trend but to a capacity cap that may be related to a level of service. We suggest that further guidance should be developed for practitioners that defines appropriate measures of reaching capacity.
- **Option B** – Guidance could be developed that involved a set of factors or multipliers which would be used to estimate benefits in the 20-60 year period based on the benefits in the 0-20 year period. Further work would be needed to determine the dimensions of these multipliers but we would expect the multipliers to account for the following:
 - the trajectory of benefits over the 20 year modelled appraisal period;
 - evidence of the saturation of demand up to and beyond year 20, with reference to the National Road Traffic Forecasts (NRTF), population projections etc.;

- capacity constraints; and
- the uncertainty about the long term future and the certainty equivalent value of the benefits.

In addition to the above, we recommend that sensitivity tests (high, central and low cases) are undertaken with appraisal, as currently is the case.

Certainty Equivalent

We draw attention to the recommendation to consider a certainty equivalent value for future benefits. Attaching more weight to a downside outcome than to a central value has the effect of reducing the weight put on the expected long term outcome and so reduces the likelihood of benefits increasing every successive year until the end of the appraisal period at year 60. This would replace the previous role that the demand cap had in accounting for the increased uncertainty of long term benefits.

1 Introduction

This Chapter provides an introduction and background to the study and introduces key concepts. It is structured as follows.

- **Overview** – provides a summary of the commission and the purpose of the work, which is to advance the discussion around demand capping in appraisal.
- **What Drives Long Term Benefits?** – introduces the concept that the primary driver of long term benefits is changes in demand and its interaction with supply via crowding and congestion feedback.
- **Demand Cap** – describes the current practice of capping demand in cost benefit analysis.
- **Limitations Of The Demand Cap** – describes issues with the current practice of capping demand in that long term benefits are not captured.
- **Uncertainty** – introduces the important concept of uncertainty.
- **Theoretical Background** – an introduction to the transport economics that underpins the simulation tests, particularly the concept of equilibrium between supply and demand.
- **Certainty Equivalence** – introduces the concept of a certainty equivalence and how it could be used in cost benefit analysis.
- **Study Approach** – describes the approach to the study and the report structure.

1.1 Overview

Arup and the Institute for Transport Studies (ITS), University of Leeds were commissioned by the Department for Transport (DfT) to undertake research into the appraisal of the long term economic benefits of transport schemes. This study builds on the work by ITS Leeds on *Specifying the Demand Cap for Rail*¹ (which was completed in April 2013) as well as other work in this field.

The scope of this study is broad and the intention is to **advance the discussion** around the demand cap and quantifying long term benefits, including recommendations for further work. The study is not intended to provide solutions to every question surrounding this topic.

This study considers the longer term quantification of ‘core benefits’ which in practice are often substantially made up of transport user travel time savings and crowding/congestion relief, which are driven by changes in the pattern of demand. Wider impacts, which are often referred to as **wider economic benefits are**

¹ The Specifying the Demand Cap for Rail Study was undertaken by John Bates, Tom Worsley, Mark Wardman, Chris Nash and John Preston (Bates et al)

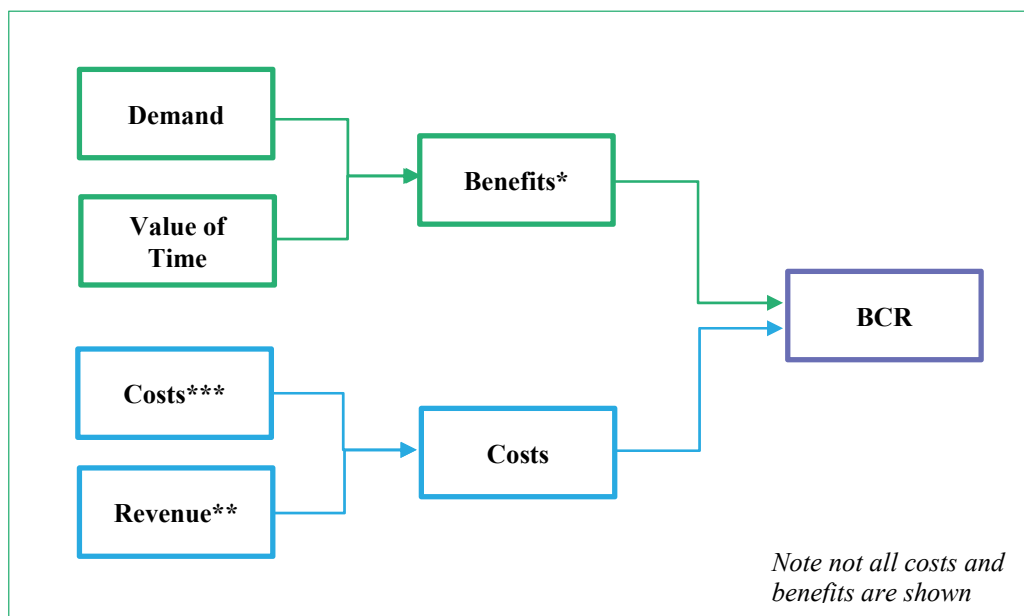
excluded from the scope of this study as are qualitatively-measured benefits related to the environment, safety and other social impacts.

1.2 What Drives Long Term Benefits?

The benefit to costs ratio for a prospective transport investment is made up of a mixture of financial and non-financial costs and benefits, which are derived by examining the change between a ‘do something’ and a ‘without scheme’ case.

Figure 1 shows a simplified diagram of the key inputs to the Benefit Cost Ratio (BCR). This diagram is deliberately simple in order to highlight the importance of the transport demand modelling in driving economic benefits.

Figure 1 – Simplified benefit-cost ratio drivers



* demand (including consideration of crowding) and value of time are used to quantify travel time savings, the primary benefit

** revenue, to public transport operators, is treated as a negative cost

*** costs include maintenance and capital costs

Demand

The economic benefits typically quantified in Cost Benefit Analysis are summarised below. These are all driven primarily by a shift in journeys / trips from longer routes to shorter routes (or from crowded/congested routes to less crowded/congested routes). This saves passengers and/or drivers time, and has benefits to non-users and the environment.

- user benefits;
 - travel time savings;
 - crowding relief;
- non-user benefits;

- decongestion;
- highway infrastructure;
- car accidents;
- local air quality;
- noise; and
- greenhouse gas emissions.

Demand is typically estimated using two methods:

1. a '**complex model**' phase, including a base year, the scheme 'opening year', and another model year 15-20 years after opening; and
2. a '**simple model**' phase, which currently involves capping demand after 20 years but it could potentially involve other simple extrapolation methods. In practice, a 'complex model' is usually not developed for periods more than 20 years after scheme opening.

Beyond the capped year the main determinants of changes in discounted benefits are the growth rate of the value of time and the discount rate.

Growth in the value of time

The value of time is an important driver in the quantification of benefits. Value of time increases in line with growth in average real GDP per person, which in the future is projected to be around 2.1% per annum. This means, that after the 'capping' period, benefits can continue to grow by 2.1% per annum.

Discounting

Another important factor in the long term quantification of economic benefits is the discount rate. Discount rates set by Treasury mean benefits occurring in later years are less valuable than benefits occurring in earlier years. Green Book guidance gives decreasing discount rates depending on the model period as shown in the table below.

A low discount rate means long term benefits are important to the BCR calculation. The low discount rate was introduced in the Green Book in 2002, with the concomitant increase in the time horizon to 60 years in WebTAG. The result is that the level of benefits predicted for later years has had much more influence on the present value of benefits (PVB), while at the same time the ability to predict these is uncertain.

Table 1 – Green Book Discount Rates

Year	Discount rate
0 – 30	3.5%
31 – 75	3.0%
76 – 125	2.5%
126 – 200	2.0%
201 – 300	1.5%
301 and over	1.0%

Source: *WebTAG Datebook: Unit 3.5.4*

1.3 Demand Cap

WebTAG guidance² (TAG unit A5.3 Section 2.3.1) recommends the use of a demand cap when quantifying benefits. Zero growth in demand is assumed beyond the cap year.

The rationale for the demand cap is that it:

- treats all schemes equally;
- is easy to implement;
- accounts for uncertainty; and
- availability of transport models.

The approach for a demand cap varies by transport mode:

- **Rail** – In rail, the growth in demand is typically³ assumed to stop (is ‘capped’) after a twenty year period. The rationale for this is based on the fact that the elasticities used are estimated with demand data over approximately the last 20 years. As these relationships might not always hold, it is currently assumed that growth stops and benefits are capped after 20 years.
- **HS2 (fixed level approach)** – Demand is assumed to increase to a fixed level rather than to a given year in the future, and the level assumed is approximately a doubling of trips for journeys greater than 100 miles. Once demand is forecast to hit this level, demand is capped. This effectively leads to a floating demand cap year, which changes as a result of changes in the forecasting method and/or economic forecasts.
- **Road** – In road projects, it is usual practice to assume demand is capped after a 15 year period from scheme opening (and given the lead times for projects this is considered comparable to the rail cap). Because of

² See TAG unit A5.3 for rail schemes.

³ The guidance also suggests that tests on a cap applied after 10 and 30 years of growth should be presented, and that under exceptional circumstances, such as with long-term infrastructure projects, it may be appropriate to use a different demand cap.

uncertainty increasing with forecast periods, it is pragmatic to use a consistent demand cap when comparing schemes.

Treats all schemes equally

One rationale for capping demand is that having a consistent approach allows for 'fairer' comparison between transport schemes. This is practically important when funds are limited and only a certain number of projects will be chosen.

Opponents argue that the cap is applied differently for different modes and is biased against schemes that do not reach capacity quickly – calling into question how 'fair' the cap actually is. For example a scheme with a gestation period to opening date of ten years is clearly treated differently than a project with a gestation period of two years.

Easy to implement

The current guidance is easy to understand and follow from a practitioner's point of view.

Accounts for uncertainty

Uncertainty around forecast inputs and parameters increases the further in time forecasts are made. In addition, many factors not captured in traditional transport models may affect rail or road demand in the long run, for example the potential rise of autonomous vehicles. The demand cap represents a pragmatic limit to control for uncertainty.

The HS2 economic and strategic cases published in October 2013 provide an example of the demand cap being used as a proxy for uncertainty. It shows the effect of changes in the year in which demand is capped on the project's BCR. The HS2 analysis was used to demonstrate the scenarios for the cap and for the exogenous demand drivers under which the high value for money category of the scheme's central case would change.

Availability of transport models

Transport models and appraisal rely on quality inputs. Many macroeconomic forecasts, used as inputs in transport models, run for ~25 years. This makes transport modelling beyond this point challenging. There are also many practical considerations about model convergence and the time and cost required to run additional transport models within a project. The demand cap effectively eliminates this problem by limiting the period by which demand needs to be modelled.

1.4 Limitations Of The Demand Cap

One limitation (but also a strength) of the demand cap is that it is a simple way of taking account of a set of complex phenomena. The rationale for imposing a cap at any fixed date such as twenty years into the appraisal period was mostly pragmatic rather than based on evidence or theory. DfT needed to respond to the extension of the appraisal period to sixty years in a context where the period for

traffic and benefit modelling was typically twenty years. Now DfT is essentially asking ‘what else could we do’ and ‘would this be better?’

There are three phenomena which lie at the heart of answering this question:

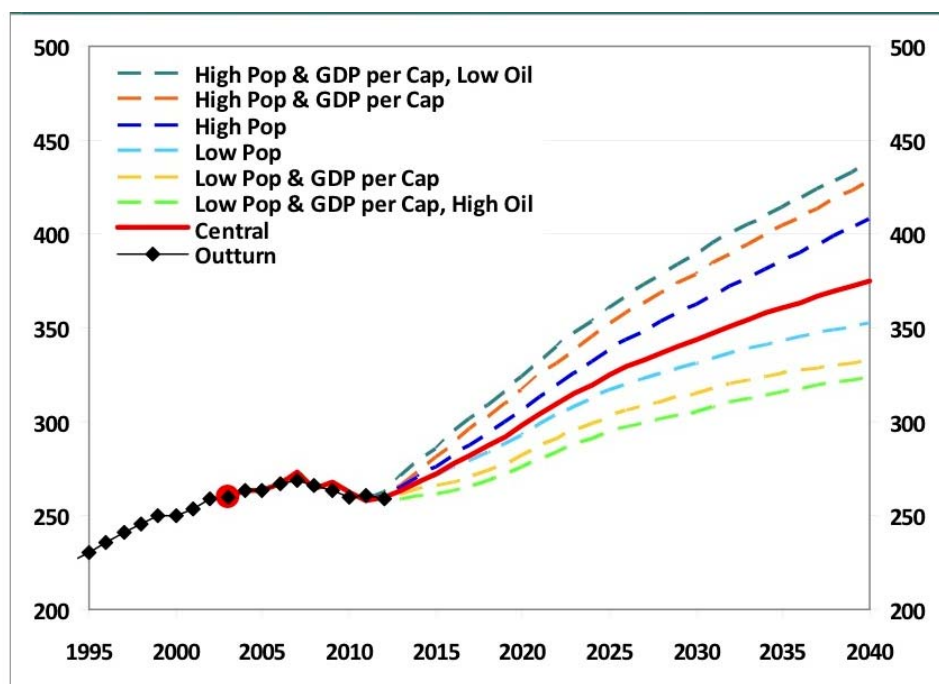
1. what is the ‘true’ long term demand of transport schemes?
2. what role does uncertainty play? and
3. how do we model capacity, crowding and congestion in the long term?

The **first** is the ‘true’ or rather ‘expected’ long term growth rate of underlying demand driven by factors such as income per capita, population and other drivers unrelated to particular schemes. Conceptually the cap could imply some form of saturation (e.g. the income elasticity of demand for travel falling to zero at some future date). This is explored further in Chapter 2 Historic Demand Trends and Chapter 3 Historic Scheme Trends.

The **second** is uncertainty. Inherent in decisions about long term transport infrastructure is the unfortunate fact that we do not know the future trajectory of travel demand and its drivers with certainty. Even if we can estimate expected (mean) growth rates, these have variance associated with their central values.

Both for traffic and for unit benefits such as travel time values DfT has illustrated this with the use of fan diagrams. The volume of traffic next year may be known to within plus or minus 2 per cent. But the volume of traffic in 60 years’ time may only be known to within plus or minus 50 per cent as illustrated by Figure 2. The question is how, in the appraisal, DfT should take this into account. Conceptually, the cap is a broad brush way of taking account of the fact we are dealing with wide ranges over the second half of the appraisal period.

Figure 2 – England Traffic forecasts



Source: Department for Transport, Road Transport Forecasts 2013

The **third** is capacity. Suppose we simply removed the cap and extrapolated traffic and benefits from year 20 to year 60 in an unconstrained manner. Would that be satisfactory? Essential to the integrity of the appraisal method is the realism of the Do-Minimum and Do-Something which are compared against each other. If schemes are predicted to fill up at some point in the appraisal period, then it is reasonable to expect the benefits growth curve to be attenuated, possibly quite sharply. The cap could be a form of insurance policy against presenting appraisal results which are unrealistic because a capacity limit has been reached and passed.

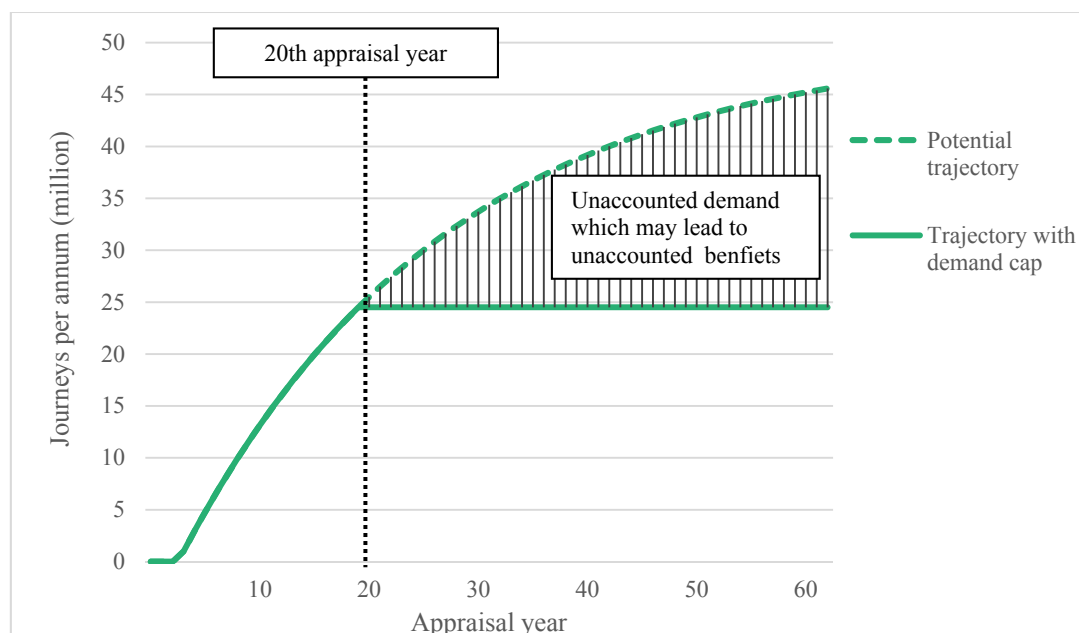
In order to investigate this problem, we have brought two approaches to bear. The first is to look for relevant evidence. The second is to undertake a desktop exercise to try to represent these phenomena in a simulation model.

Illustration of the limitations of capping

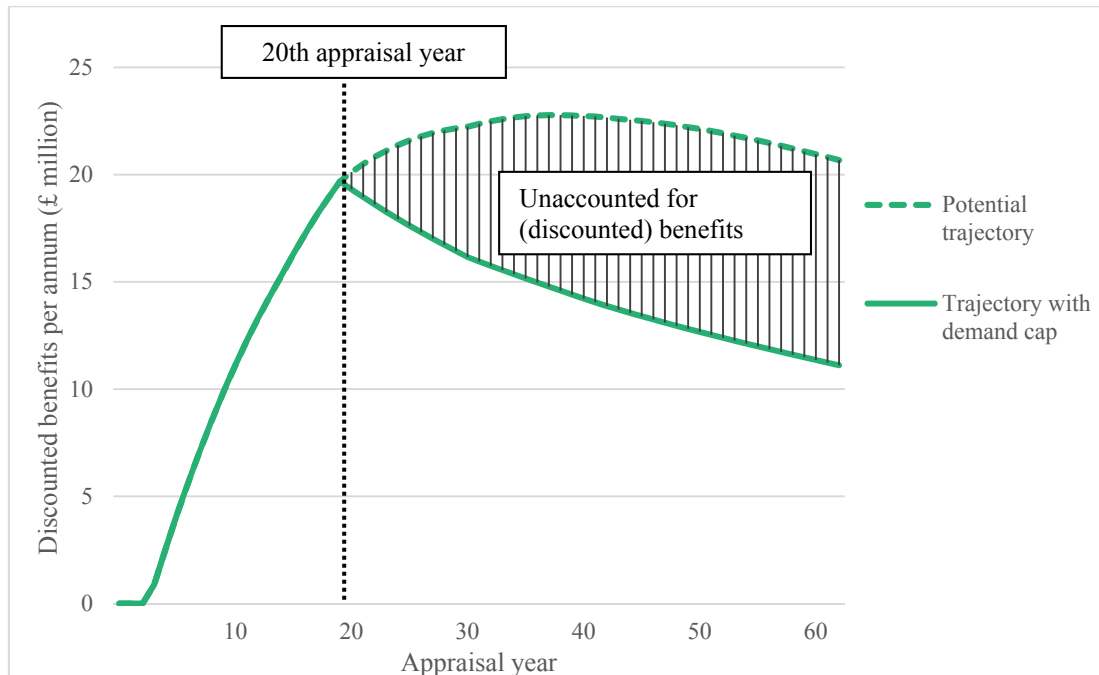
The problem with the demand can be illustrated in the following way. First, if everything goes according to plan, and leaving aside capacity considerations, capping causes us to miss the benefits associated with demand growth. **Figure 3** illustrates this by showing how potential growth in demand after the 20th appraisal year is unaccounted for when the cap is applied, potentially reducing benefits (see **Figure 4**).

Second, we do not know the future, so throughout the appraisal period we are dealing with a fan of outcomes rather than a single valued distribution. Thirdly, the chances are that for most schemes we will run into capacity problems during the appraisal period with consequences for scheme benefits once capacity is reached.

Figure 3 – Demand profile for a hypothetical rail scheme



Source: Arup (illustrative example)

Figure 4 – Discounted benefits profile for a hypothetical* scheme

Source: Arup (illustrative example)

* This is a hypothetical example to illustrate the point that there is a difference in benefits if demand is capped or uncapped

Alternatives to the demand cap

This study explores potential alternatives to the demand cap, including applying the demand cap further in the appraisal period (e.g. 30 or 40 years), using a floating cap based on a demand threshold (as previously done for HS2), or developing a ‘certainty equivalent’ value for benefits to account for issues with uncertainty. The following section provides an introduction to the ‘certainty equivalent’ value.

1.5 Uncertainty

Uncertainty can be defined as inevitable ‘unknowns’ that arise as part of appraisal, upon which a statistically derived understanding of probability cannot be achieved. Although the sophistication of modelling approaches has increased dramatically in recent years, scheme costs and benefits cannot be assessed indefinitely. Uncertainty arises due to a lack of ability to predict and accurately forecast demand and how this will tie with future economic, social and political conditions, along with individual behaviours and preferences, all of which will have a significant impact on the usage and vitality of our national infrastructure.

The WebTAG guidance on uncertainty in project appraisal states that forecasting uncertainty can be divided into five primary categories as follows.

- **Errors in Model Parameters** – a reasonable level of uncertainty has to be assumed due to potential errors in the parameters used for building forecasting models.

- **National uncertainty in travel demand** – uncertainty in the use of demographic projections as well as changes in consumer travel behaviour and preferences.
- **National uncertainty in travel cost** – uncertainty regarding economic variables (for example fuel cost) and government policy.
- **Local uncertainty in travel demand (scheme vicinity)** – uncertainty regarding local developments that are likely to impact on future demand for travel (i.e. housing, office space, retail).
- **Local uncertainty in travel supply (scheme vicinity)** – uncertainty regarding other local transport projects near the scheme.

One way that uncertainty can be accounted for in appraisal is through the use of a certainty equivalence which is further explain in section 1.7 Certainty Equivalence. Macro variations in transport demand is explored further in Chapter 2.

1.6 Theoretical Background

In this section we introduce the transport economics theory that underpins the Simulation Model and tests described in Chapter 4. A key concept is that of the equilibrium between demand and supply. This requires an understanding of how the demand curve shifts over time, and how its interaction with the supply curves leads to variations in the level of benefits.

In addition, we discuss the current practice of inter-/extrapolation, and changes in the values attributed to different components of benefit.

1.6.1 Discounting Benefits

When conducting appraisals, governments (or private sector investors) need to take account of how costs, revenues and benefits are distributed over time, for at least the following reasons:

1. prices change over time;
2. resources used to produce a future benefit could be invested elsewhere;
3. diminishing marginal utility of income;
4. whilst most people care about future populations, they may care less about distantly future populations; and
5. there is a chance that future benefit or cost may not occur.

Points 2 to 5 are normally handled by the practice of ‘discounting’. Point 1) is technically undisputed, but generally handled by forecasts of expected price or cost changes, fuel prices being the best example in transport appraisal.

According to the practice of discounting, if a scheme has an economic life of Y years, then the cumulative annual benefits can be standardised (or ‘discounted’) to a given base year, such that points 2) to 5) are appropriately accounted for.

Having discounted the benefits in this way, we refer to the cumulative benefits as the ‘present value’ of the benefits in the base year.

More formally, if the benefits in year Y are written as B^Y , then the present value of benefits (PVB) is given as

$$PVB = \sum_Y \frac{B^Y}{(1+r)^{Y-W}} \quad (1)$$

where r is the discount rate⁴, and W is the base year. These benefits will be defined in stated constant price units for a defined year (which need not be the same as the “base year”). From now on we ignore the discounting and focus on the undiscounted quantity B^Y .

1.6.2 Deriving Transport User Benefits

The standard way in which benefits are calculated is by means of a transport model, as we discuss below. Ideally the model would be run for every year Y of the benefit stream, but in practice this is not generally feasible. Below we discuss the assumptions required about the path of benefits for those years when explicit model runs are not carried out.

Within the benefits B^Y , we focus on transport user benefits S , and ignore distinctions of purpose and other possible “segmentations”, though in practice they will need to be made.

In terms of transport user benefits, there are a number of generalised cost components that need to be distinguished in order to identify and estimate the benefits of the transport scheme. In particular the TEE table identifies the following items: travel time, vehicle operating costs, user charges (including fares) and operator revenues: we notate these as k . According to the rule-of-a-half convention recommended by WebTAG, the contribution of component k to overall transport user benefits in year Y is given as:

$$S_k^Y = -\frac{1}{2} \sum_{ij} (T_{ij}^{\prime Y} + T_{ij}^Y) (C_{ijk}^{\prime Y} - C_{ijk}^Y) \quad (2)$$

where T is demand, C is (generalised) cost in money terms, and the prime (') denotes the “after” (with scheme) case. i and j are zones (in practice the summation can be extended over modes and time periods).

⁴ The discount rate is specified in Table A1.1.1 of the TAG Data Book, based on The Green Book. The discount rate is assumed to fall for very long periods because of uncertainty about the future.

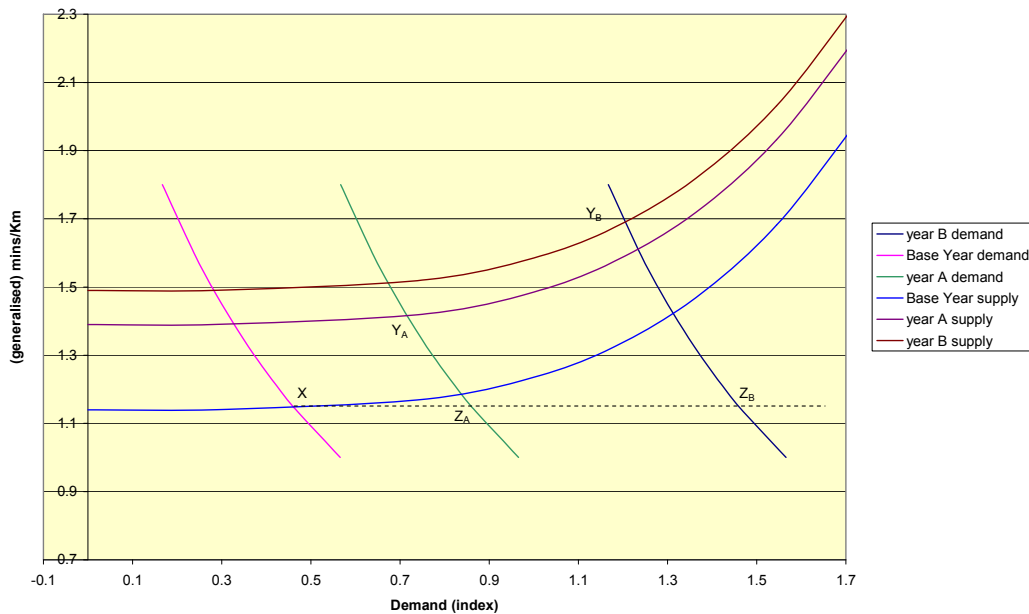
Running the transport model in years A and B allows the calculation of the quantity S_k^Y . It is clear from the definition in Eq (2) that it requires both demand estimates (T) and (generalised) cost estimates C.

The demand estimates are determined by factors, some of which are exogenous to the transport model (through population changes and income changes, the latter also affecting car ownership), and some, such as the fuel and time costs of using the scheme, are endogenous, due to supply-side effects.

1.6.3 Forecasting the effect of exogenous changes on demand and supply

In the next sections our intention is to develop a generic supply/demand simulation model which has properties in line with general WebTAG recommendations. In the base year 0 we have an observed demand T_0 and cost C_0 represented by point X. The base year demand and supply curves are constructed to intersect at this point. The properties of these curves (i.e. shape and location) are dictated by the minimum levels of demand/supply in combination with the endogenous factors detailed above, for given exogenous and scheme conditions.

Figure 5 – Illustration of Supply-Demand equilibrium over time



Source: ITS

The exogenous increases in demand to the modelled future years A and B, which entail shifts in the demand function (parallel to the Y-axis), are then predicted on the assumption that (generalised) costs remain at C_0 (as shown by the dashed line). This gives the points Z_A and Z_B in the figure, and allows us to locate the future demand curves.

However, this is merely the “latent” growth in demand: to obtain the actual demand we have to find the point of equilibrium with the future supply curves. Even if there is no change in capacity, there is still the possibility of exogenous

changes in (real) costs: these could be directly related to prices (e.g. fuel prices, fares) or to changes in the value of time. They are reflected in the shifts (parallel to the X-axis) in the supply curves for the future years, giving rise to the new points Y_A and Y_B . If there is no change in real input prices over time (fuel cost, travel time), the supply curve will not shift except as a result of a change in capacity due to a scheme, a simplifying assumption we make from now on.

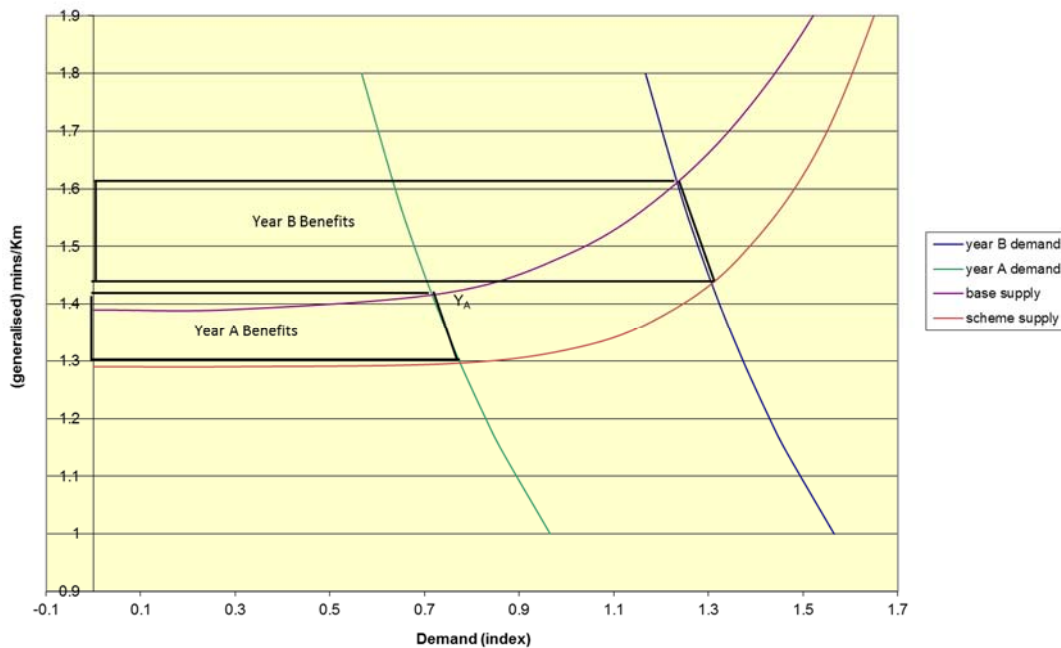
Purely for the purposes of illustration, we have drawn the curves so that the growth in latent demand is greater between the two forecast years than between the base and first forecast year, while the converse is true for the growth in costs. The result is that the growth in demand from point X to point Y_A is less than the growth from Y_A to Y_B .

1.6.4 Forecasting the effect of the scheme on demand and supply

A similar analysis is required to deal with the change in the supply curve resulting from the scheme, which typically entails a shift in supply (reflecting an extension in capacity). This then produces the quantities (T^Y_{ij} , TY_{ij} , C^Y_{ijk} , CY_{ijk}) and allows the benefit component S_k^Y to be computed as in Eq (2).

To illustrate this, we return to **Figure 5**, but concentrate on the two future years, A and B. To simplify matters, we assume that the supply curves do not change over time (implying, among other things, that fuel prices, fares etc. remain constant in real terms).

The previous “Year A Supply curve” is now assumed to represent the base case in both years A and B, and to this we now add the scheme supply curve. On the assumption that this increases capacity and reduces travel time, this lies below the base supply curve, as shown in **Figure 6**.

Figure 6 – Illustration of Benefits over time

Source: ITS

In this figure, the benefit trapeziums in the two years are enclosed in bold black lines. Of course, the curves are all merely illustrative. But in this case, the benefits in year B are considerably greater than in year A. This can be seen to arise, firstly, from the overall (horizontal) shift in the demand curve (generally reflecting increased income and population), and secondly from the increased cost savings (vertical) in the later year, reflecting the greater congestion with the Do-Minimum as a result of increased demand.

Of course, these cost savings will start to decrease once significant congestion is experienced on both base and scheme networks. But they are unlikely to decrease sufficiently to outweigh the increase associated with demand.

This is the general case. The relative sizes of the trapeziums depend on the supply and demand elasticities (i.e. reflecting the slope of the functions) and the starting and final positions of demand relative to supply (i.e. reflecting the location of the functions), and these will vary from scheme to scheme.

For example, if we consider a rural road upgrade, where demand is well below capacity (so that we are on the horizontal part of the supply curve), the height of the trapezium will be fixed (except perhaps in the very long term), and the benefits will grow in line with the shift of the demand curve. But for a congested urban scheme, we are in the position of the “year B” case in Figure 2, where the height is highly dependent on the location of the equilibrium point on the steep upward-sloping part of the DM supply curve.

To recap, the location of the demand curve in year Y will be dependent on demographic and income effects, particularly those relating to car ownership (as well as, for uni-modal models, changes in the costs of alternative modes), while

the supply curves for the base (“Do-Minimum”) and “scheme” cases will be dependent on capacity and exogenous costs.

1.7 Certainty Equivalence

This section describes the concept of the ‘certainty equivalence’ which is referenced later in this report. In essence, applying a certainty equivalence means that more weight is placed on the downside outcomes in demand and benefits.

Following Treasury/WebTAG rules, the discount rate i is considered to be risk free, and a reflection of pure time preference – in other words, indifference between £1 now and £(1+i) next year. However, the implications are that the quantity to be discounted should be the ‘certainty equivalent’. This is related to the risk premium which is the amount by which the risky asset’s expected return must exceed the risk-free return in order to make the risky and risk-free assets equally attractive (assuming risk averseness).

In standard theory, the certainty equivalent is defined as the guaranteed amount X_c whose utility is equal to the expected utility of the uncertain outcome X . If X has variance σ^2 , then

$$U(X_c) = E[U(X|\sigma^2)]$$

The risk premium π is the difference between the expected outcome and the certainty equivalent, so that

$$\pi = E[X|\sigma^2] - X_c$$

For risk averseness $\pi > 0$ and is an increasing function of σ . More generally, we can write

$$X_c = E[X] / (1 + \rho(\sigma))$$

where ρ , a function of the variance σ , > 0 .

This suggests that we should discount the expected outcome, after taking account of all the factors affecting forecast uncertainty, by $(1 + \rho)$, to get the certainty equivalent which can then be further discounted to allow for time preference.

In general we expect σ to increase with time as the future becomes more uncertain (as well as whether the model parameters will stay constant), so that ρ will also increase over time. The effect can then be approximated by using the “risk-adjusted discount rate” – adding an appropriate amount to the risk-free discount rate i .

1.8 Study Approach

In order to address the three questions summarised in section 1.4, Limitations Of The Demand Cap, the study considers four key areas, which each have a chapter dedicated to them.

- **Historic Demand Trends** (Chapter 2). Given that economic benefits are quantified primarily by looking at changes in patterns of demand, and that the appraisal period is 60 years, our first task was to investigate long term trends in travel demand. The overarching question here, is what happens to travel demand in the long term, and can the past tell us anything about whether we can forecast 60 years into the future?
- **Historic Scheme Trends** (Chapter 3). This Chapter looks at travel demand for a number of individual schemes in order to see what happens to demand after 20 years.
- **Simulation Model** (Chapter 4). This Chapter considers the interaction of supply and demand from first principles and the impact that capping demand has on the equilibrium point, benefits and ultimately, the BCR. It also considers a real project case study for road.
- **Long Term Demand Modelling** (Chapter 5). This Chapter explores the practicalities associated with quantifying long term benefits, which primarily relates to long term demand forecasting. It also considers many implications and issues around long term benefits.

This is followed by Chapter 6 that summarises a **Discussion** and Chapter 7 **Recommendations**.

2 Historic Demand Trends

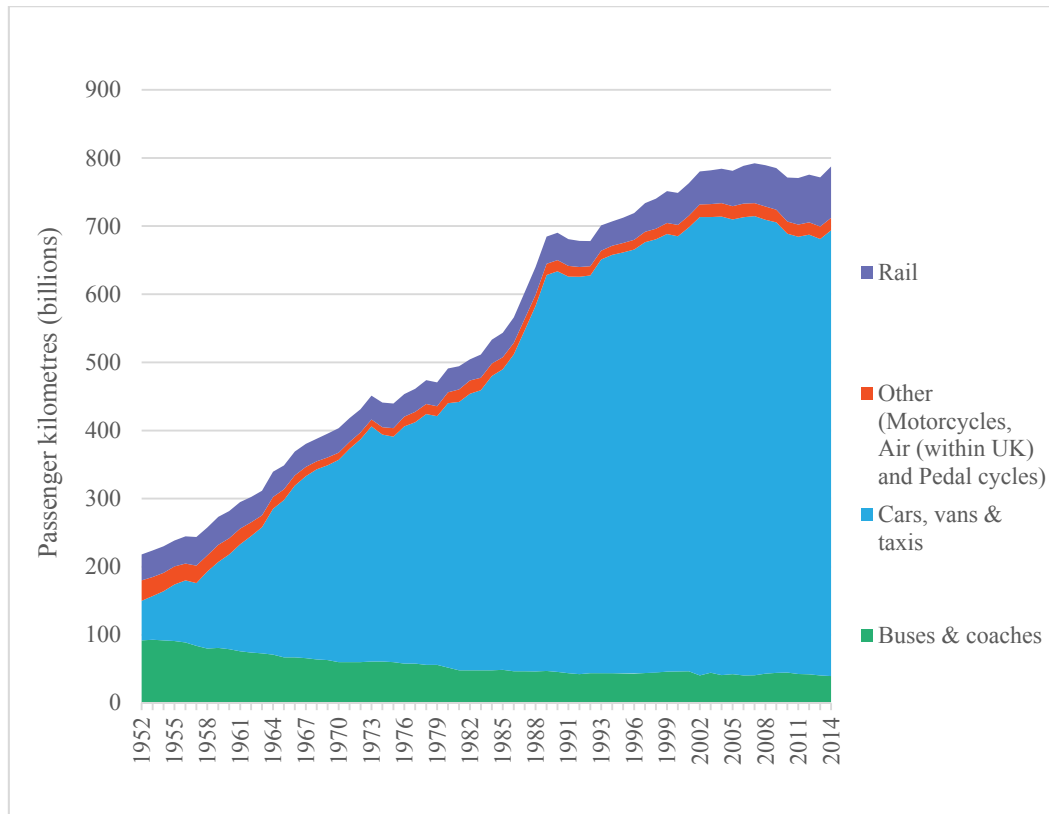
As long term demand forecasting is an important aspect of quantifying long term benefits, this chapter considers historical trends across transport, and a more specific review of rail trends.

- **Transport** – this section gives a high level introduction to long term trends in travel demand across all transport modes. It shows an overall steadily increasing growth in travel demand in Britain since 1952, with small reductions during recessions. Within this overall trend, some modes have seen their share increase (for example rail), while others have seen a reduction (bus and coach).
- **Rail** – this section takes a more detailed look at long term trends in rail demand. Of particular interest is the variation in rail journeys/trips over the ‘long long term’ since 1830.
- **Rail Market Saturation** – the key finding here is that there is no evidence to suggest that there is saturation in the rail market.

2.1 Transport

This section considers trends in transport overall. **Figure 7** shows trends in travel demand by mode since 1952. The key points are as follows.

- Overall there has been a steady increase in passenger kilometres travelled.
- Kilometres travelled by car, van and taxis increased between the 1950s and the 1990s, save for mild declines at the time of the 1970s oil crisis and the early 1990s recession.
- Since the 1990s trends have been more mixed, initially leading to a suspicion that **peak car** had been reached, but in very recent years there has been an upwards trend that has been attributed to the falling price of oil, improvements in vehicle fuel efficiency and the rising use of light goods vehicles, particularly for deliveries.
- Kilometres travelled by **rail** remain a small proportion of the sample (10% in 2014), despite the absolute increase over time, with 2014 seeing the highest kilometres travelled since the Second World War. Given that rail travel makes up a small proportion of travel overall, *small changes in rail-road mode share can lead to big increases in rail journeys*. This in part explains the larger percentages changes in rail journeys in recent years.
- Buses and coaches have declined steadily since the 1950s, with the steadying in most recent years explained by growth in London that has been counterbalanced by decline elsewhere.

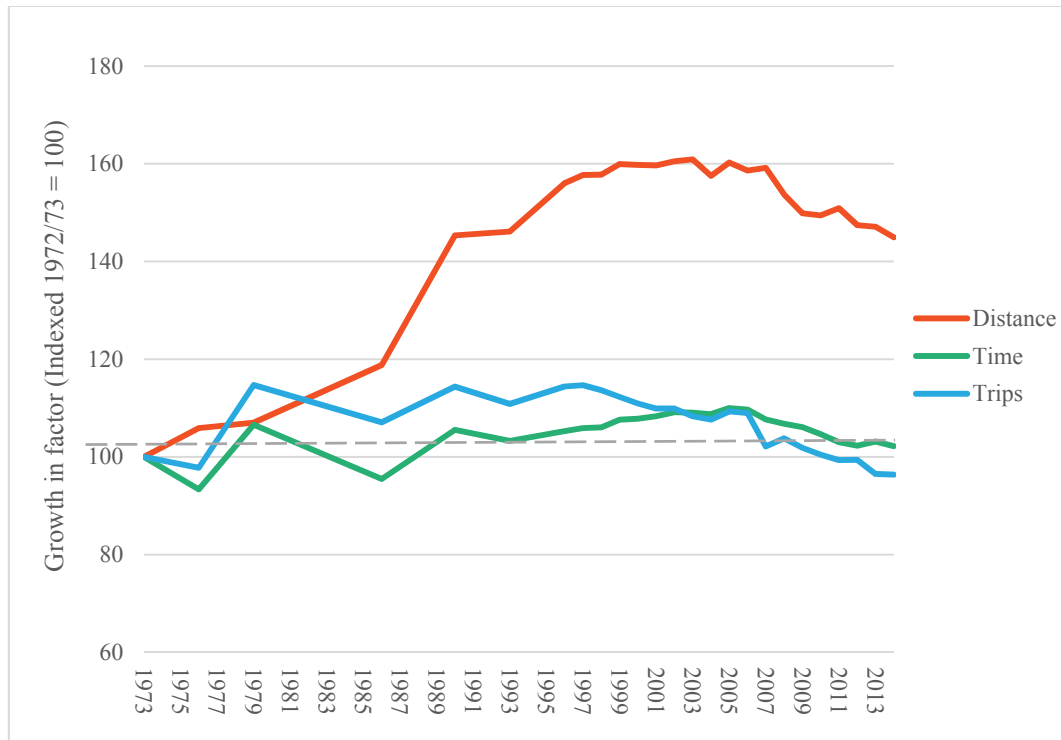
Figure 7 – Passenger kilometres by mode in Great Britain, 1952-2014

Source: DfT – Table TSGB0101

The average number of trips and the time spent travelling has seen little change since 1972, as shown in **Figure 8**. The figure also shows:

- the **distance travelled has increased significantly**, reflecting the increased average speed resulting from the completion of the motorway network, and the increased popularity of higher speed modes such as intercity railways (see following section) and the falling cost and greater availability of aviation.
- **From 2007 onwards there has been a decrease in distance travelled, time spent traveling and number of trips taken, compared with previous years.** These downward trends have perhaps been caused by the recent recession, or by the continued substitution of the internet for forms of leisure activity (the greatest declines in trip rates, in recent years, have been in visiting friends at home, sports and family and shopping).
- The **internet** is also likely to have facilitated the growth in the popularity of home working during the mid-2000s (which has now plateaued), contributing to a fall in commuting trips.

Figure 8 – Trends in trips, distance and time spent travelling by a typical person – indexed



Source: DfT – Table NTS01015

The recent, continued decline in both trip rates and distance suggests that there could be a potential saturation of overall transport demand. Together with the above, this suggests that there is an interaction between the prevailing economic conditions, the pricing and availability of competitive modes, technology, demography and other factors⁶ in determining transport demand.

Models for forecasting car ownership and road traffic have, arguably, been more stable with much of the cause of forecasts differing from outturn being explained by the main drivers of demand differing from those assumed when the forecasts were made:

- the growth in the number of households;
- in household income; and
- at least on a cyclical basis, in fuel costs.

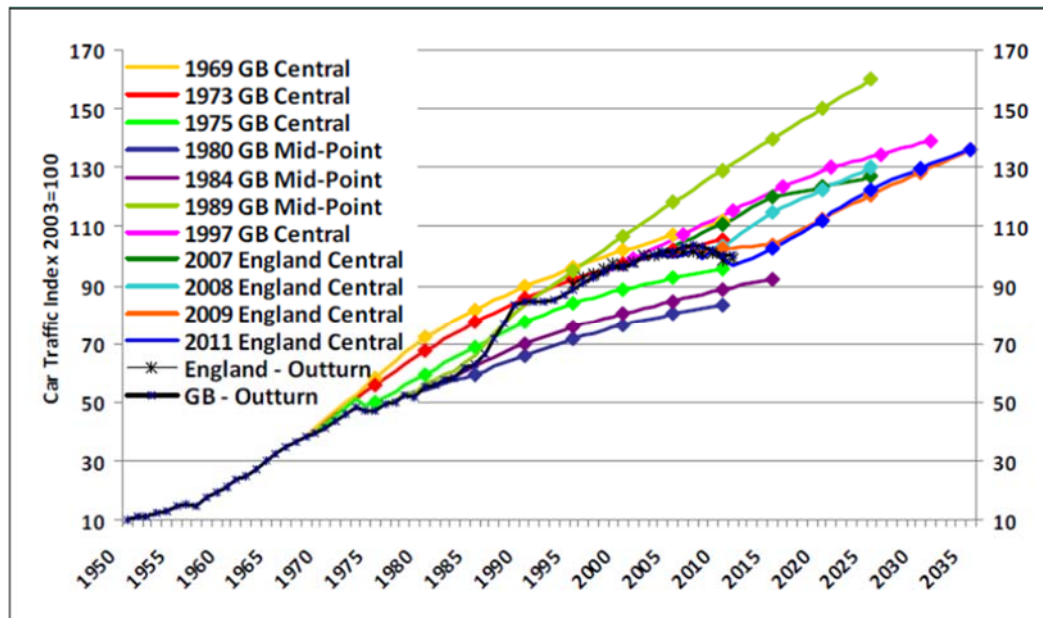
Figure 9 shows the DfT’s forecasts of road traffic dating back to the 1969 projections, based on the logistic curve aggregate model developed by John

⁵ Figures prior to 1989 are for Great Britain, rather than England only. Figures prior to 1995 are based on unweighted data.

⁶ For example, recent work for the London Roads Task Force suggested many households did not own a car due to “the stress of owning and driving a car in London” and the proximity of local services. Cultural factors such as the absence of a positive social stigma for car ownership among young people has also been cited as an explanatory factor in the fall in the rate of increase in car ownership.

Tanner and described in the TRRL report LR650 (1974). Moreover, potential causes of differences between forecast and actual trends in road traffic are more open to analysis than is the case of rail forecasts as was demonstrated by the analysis described in the National Road Traffic Forecasts 2015.

Figure 9 – DfT central or mid-point car traffic projections and outturn statistics



Source: DfT

2.2 Rail

This section provides a summary of historic rail demand. It considers demand in terms of two main measures which are published:

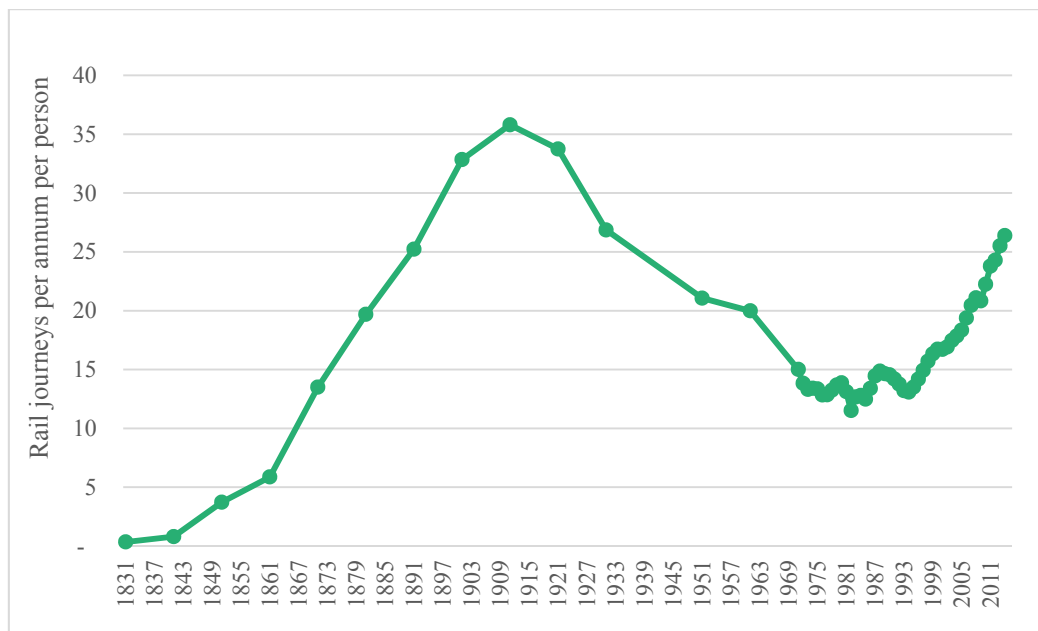
- data on **passenger journeys**, derived from the number of tickets sold, which includes estimates of the usage of season and other multiple journey tickets; and
- data on **passenger kilometres** travelled, again derived from the ticket sales data.

On a per person basis rail has, broadly speaking, seen two turning points in the last 180 years, as shown in **Figure 10**.

- Firstly, there was the *growth in popularity leading up to the First World War*, at which time the rail network was more geographically extensive, and rail was perhaps the only mode of transport for journeys longer than 5-10 miles that was accessible to most of the population.
- Then, the steady decline in popularity that lasted for most of the twentieth century, until the mixed fortunes of the 1980s and the upturn at the time of privatisation.
- Rail journeys on a per person basis are currently increasing, but have some way to reach the high point in 1911.

A reversal of this appears unlikely (although possible) under the current industry structure. Even the Hatfield effect and its aftermath in 2000/1, considered by many to be the low point of the privatised rail industry, produced only a small inflection in the upwards trend.

Figure 10 – Estimate of rail trips per person p.a. in Great Britain, 1831 to 2014



(Note: Graph plots National Rail journeys divided by GB population in that year.)

Source: ATOC, DfT, ONS and Whitaker's Almanack⁷

Furthermore, a future turning point that is driven by technology should not be ruled out. Fully autonomous vehicles have the potential to allow car drivers to use time productively (matching one of the key competitive strengths of rail travel) and new ownership models remove the need to find a parking space (matching another). The impact of these future technologies depends as much on legal matters as technological ones, including whether drivers would be required to pay attention to the road, in the same way that an airline pilot has to when applying autopilot.

Passenger kilometres remained remarkably constant between the early 1950s and the mid-1990s, with some decline over the 1960s. However, passenger journeys showed a different pattern, declining gradually up to the mid-1970s as average trip lengths increased, a pattern which is accentuated when population growth is taken into account to show trends in rail trips per person.

Before the mid-90s route length broadly matched passenger journeys (as routes were closed as demand reduced). However in recent years the network has been used more intensively, and passenger journeys have surged, as on-route upgrades such as frequency and train lengths have in many cases been outstripped by demand, leading to an increase in average loadings. With the exception of the

⁷ Passenger journeys 1830 to 1949 are from ATOC. Passenger journeys 1950 – 2014 are from DfT –TSGB0102. Population 1951 to 2014 is based on ONS data. Population from 1831 to 1931 is based on data from Whitaker's Almanack.

period following the most recent financial crisis, journeys per person have increased year on year since 1994.

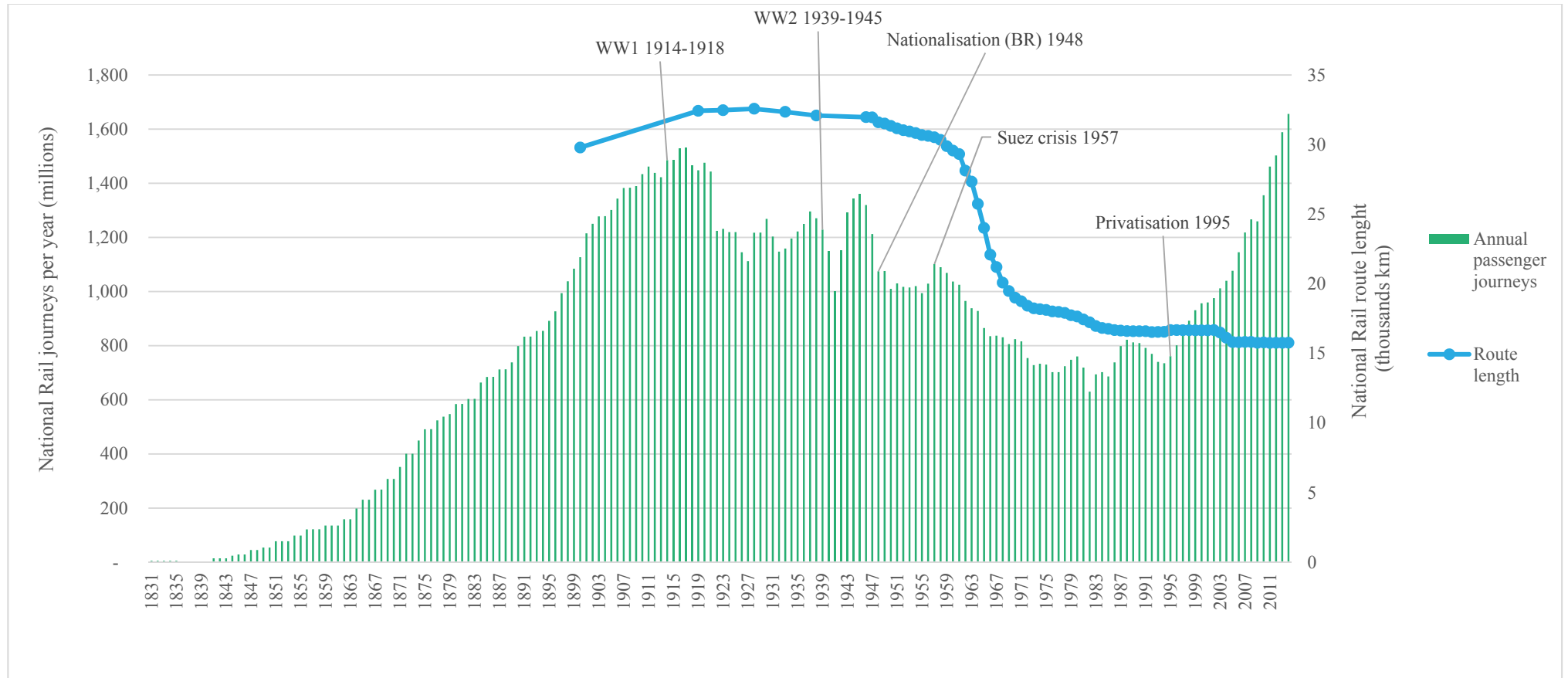
While much of the analysis underpinning rail forecasting models was published as academic research, details of the models themselves have not been made public. PDFC, the Passenger Demand Forecasting Council, managed by ATOC, took over responsibility for the demand forecasting handbook (PDFH) on privatisation. Research into updating and improving the methods is funded by members of PDFC and the handbook is only available to members.

The British Rail passenger forecasts, dated June 1982, made no explicit allowance for long term relationship between income growth and rail demand. The forecasting model includes fare and journey time elasticities and a method for converting frequency changes into changes in journey times. The journey time elasticities were set so as to be consistent with the fares elasticities and with the relative components of fares and generalised journey time with the fare elasticities. Positive changes in 'image', referred to as the 'nose cone effect' were assumed to add a one-off continuing increase in demand, based on the evidence of the introduction of the HST fleet.

All sectors – Inter-City, London and the South East (L&SE) and what was then denoted Provincial services – were assumed to grow by an exogenous trend factor of 0.8% p.a. for Inter-City and Provincial with L&SE having a 1.2% factor for non-commuting, for which a value of 0 was applied. In addition, the model included a cyclical economic indicator. This served to explain the decline in rail demand during a recession (there being no GDP term in the model) which picked up during the upturn until steady state GDP growth was reached, at which point rail demand reverted to trend. The 1982 report justifies this approach on the grounds that, during the recessions in the 1960s and 1970s, reductions in income growth were not associated with reductions in car ownership as car owners retained their vehicles but reduced discretionary spending on rail. With the recovery, there was initially no growth in car ownership and discretionary income was spent on rail, until the economy returned to trend and growth in car ownership resumed.

The 1986 Passenger Demand Forecasting Handbook adopted a model which is comparable to present methods and included a GDP elasticity in addition to elasticities for fares and generalised time, which was composed of in vehicle time, interchange, frequency, station access, rolling stock quality and reliability. Offsetting the effect of GDP growth was a negative time trend, with the coefficients on the GDP term (2.5% p.a. for all flows other than L&SE commuting, set at 0) generally offsetting the GDP elasticity. The report notes that with trend economic growth, any change in demand will be accounted for by changes in fares or generalised journey time. The 1986 elasticities were derived from a study commissioned by British Rail which was subsequently published (Owen and Phillips, 1987).

Figure 11 – Rail journeys and length of National Rail route in Great Britain, 1831 to 2014



(Note: Route length data before 1900 was not available.)

Source: ATOC & DfT8,

⁸ Passenger journeys 1830 to 1949 are from ATOC. Passenger journeys 1950 – 2014 are from DfT – TSGB0102. Length of National Rail route is from DfT – TSGB0601.

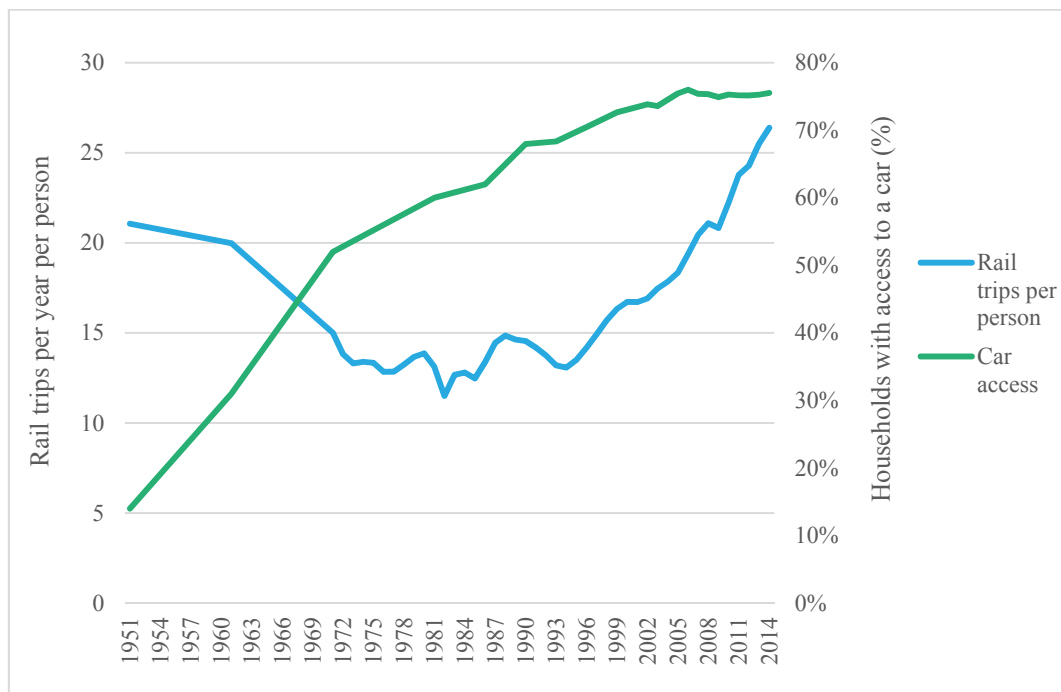
PDFH 3.1 followed the methods of the 1986 report with some upward revision both to several of the GDP elasticities and an increase in some of the time trend terms. This included income growth, offset by a negative time trend (which was taken as a proxy for increasing car ownership), and by increases in average fares.

Trends in central London employment were deemed to be the main determinant of commuter travel in London and the South East. Most of the investment approved during that period was aimed at reducing the long run costs of operating and maintaining the existing level of capacity. Improvements to service, such as the inclusion of air conditioning in rolling stock (such as the class 158 DMUs in the north of England), were required to generate sufficient additional revenue so as to fund the incremental costs.

The time trend served to reconcile the impact of increasing GDP per capita on the demand to travel by all modes (which increased over three-fold between 1952 and 1989) with the effects of factors that mitigated against growth in rail patronage, including increasing competition from other modes and changes in land use. Car traffic increased tenfold between 1952 and 1989. It was, it would seem, infeasible to construct an index which reflected the increase in car ownership, the improvements in both vehicle and road quality during this period and which might explain past rail demand and be used for forecasting.

Figure 12 shows how rail trips per person declined as car access increased, from 1951 through to the mid-80s. However, this relationship changed from the early 90s when rail trips per person began to increase, even while car access continued to increase (although it has plateaued over the last 10 years).

Figure 12 – Rail trips and car ownership



Note: Graph plots National Rail journeys divided by GB population in that year.

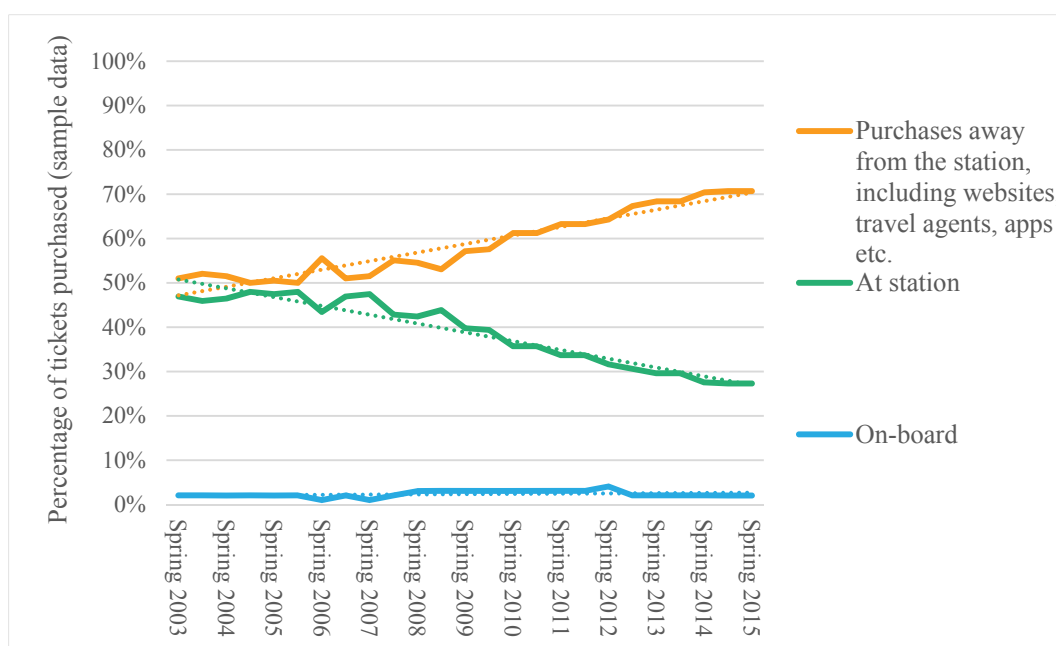
Source: ATOC, DfT, ONS

The main purpose of the earlier versions of PDFH was to provide the nationalised British Rail with evidence on which to base proposals for incremental investment in service quality. The use of cost benefit analysis under the British Rail regime was limited to the assessment of rail closures, to which ministers would rarely assent despite the strength of the case, to the Channel Tunnel Rail Link (1993) and to the schemes that were considered in the Central London Rail Study (1989). It was only in the case of London schemes that capacity constraints were perceived as an issue because of a forecast of employment growth in the central area. All other investment was justified on the estimates of the impact of the increment on demand and hence on revenues and on whether the incremental revenues would exceed the costs, and so the scrutiny to which aggregate demand forecasts were subjected to rather less than might now be the case.

The main change to the rail forecasting model occurred with PDFH 4, which quoted elasticities of rail demand for changes to other modes in addition to GDP, employment (for commuting demand), population and dropped the time trend. The range of endogenous variables, described a specific attributed, is also increased. Subsequent versions of PDFH have followed the same approach.

Several factors – perhaps not all of them fully captured by PDFH – have contributed to the rapid growth in rail demand since 1995, growth that was not anticipated at the time of rail privatisation (1994-97). Improvements on the supply side, with the privatised industry incentivised to improve the quality of service, operate more frequent trains and manage yields through fares differentiation more effectively than its nationalised predecessor. The spread of the internet also led to wider availability and wider knowledge of rail ticketing products, which also allowed passengers insight into prices before making a purchase, and removed the need in many cases to pay a separate visit to a rail station, or a call, to make a purchase, for many journeys (see **Figure 13**).

Figure 13 – Approximate change in ticket purchasing locations over time for trips on long distance rail operators, 2003-2015

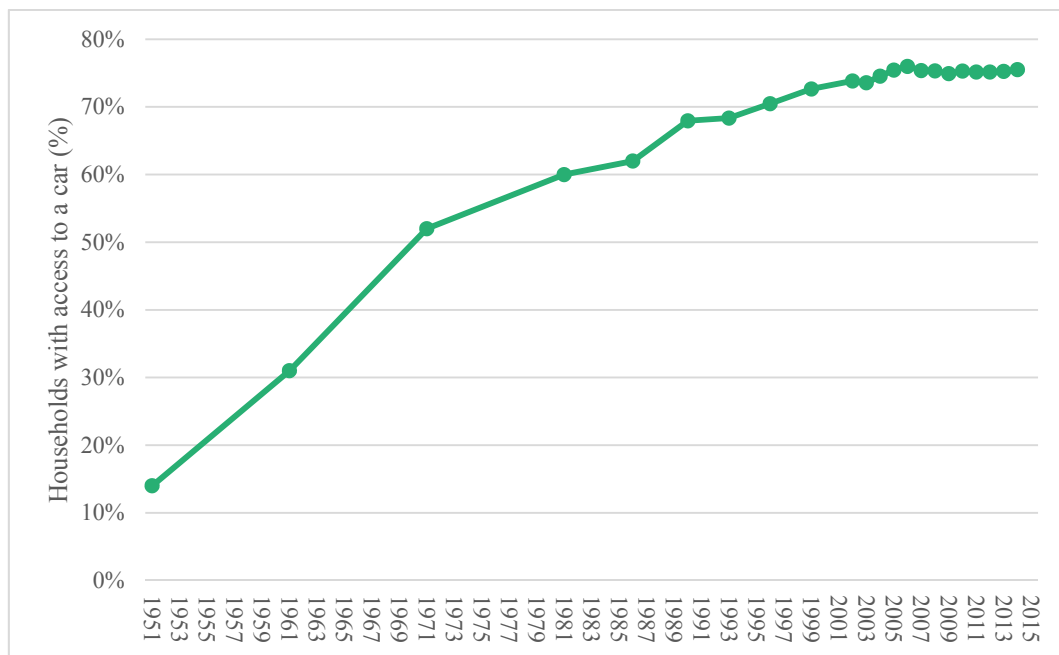


Source: National Rail Passenger Survey⁹

Demographic and spatial changes, which may only partially be accounted for by PDFH, may also explain some of the additional growth since the mid-1990s, with larger cities well served by rail becoming preferred locations for many employers and employees, aided by changes in the structure of economic activity away from locations which were less easily accessed by rail. Growth in household incomes and in GDP that were identified as key determinants of historic demand must also have driven much of the forecast growth.

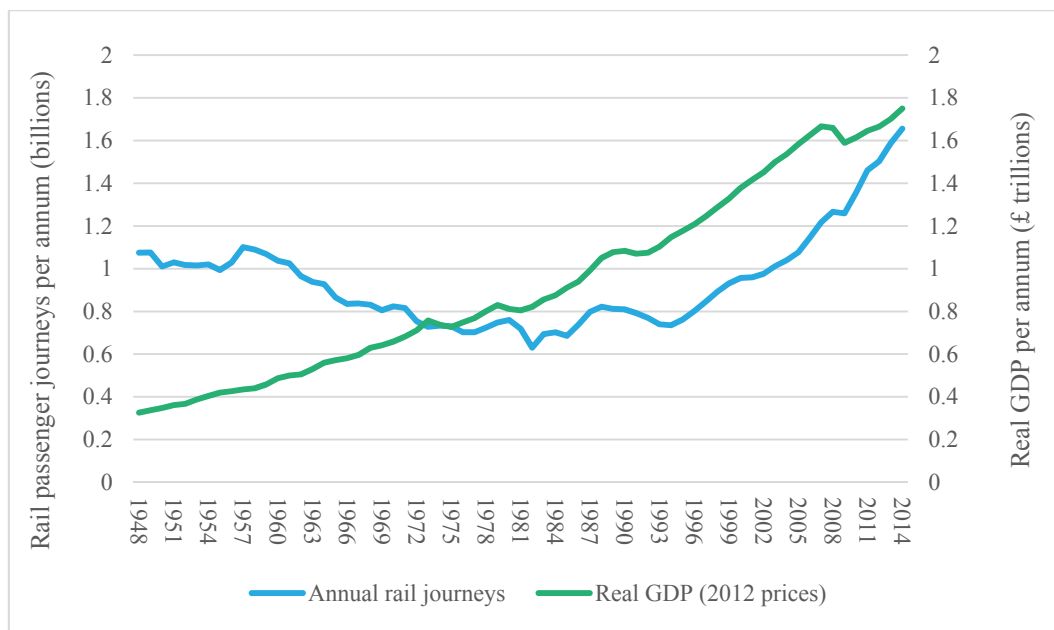
However, rail patronage continued to grow, albeit at a slower rate, during the recession when demand for other modes fell. While other factors, such as service quality improvements contributed to growth during the recession, failure of demand to decline in response to the reduction in household income calls into question the relationships on which the demand model is based. In addition, rail’s relative performance against other modes improved, with increasing congestion, a slowing down of car ownership growth (see **Figure 14**) and policies such as eliminating the tax advantages of owning a company car all had provided an impetus to rail patronage.

Figure 14 – Percentage of households in England with access to a car, 1951 to 2014



Source: DfT – NTS0205

⁹ 4,000-6,000 people surveyed depending on year. The possible answers to this question survey question have changed over time. The trends shown are based on the questions that are consistent across all years.

Figure 15 – Relationship between rail passenger journeys and GDP

Source: DfT & ONS

One observation from this brief review of rail forecasting methods is the finding that, over the past 30 years, the published GDP elasticity values for major segments have all been above unity. Over the same period, rail demand relative to GDP appears to have increased (broadly, depending on the time frame chosen – see **Figure 15**) on a 1:1 basis with GDP – at least until the 2008 financial crisis. Earlier versions of the PDFH model may have encountered problems in separating the time trends from the GDP elasticity – but the subsequent analysis which took fuller account of competition from other modes did not result in significant changes to the elasticity values might indicate a stable relationship. However, it is difficult to reconcile this conclusion with the growth in rail demand during the recent recession. And prior to the 1990s growth period, rail had a negative correlation with GDP.

2.3 Rail Market Saturation

This section summarises our view on rail market saturation. In summary, we believe *there is no evidence to suggest saturation in the rail market*.

As products reach maturity through their natural lifecycle, a phenomenon commonly observed is that of ‘saturation’. The standard representation of this phenomenon is in the form of a sigmoid growth curve, a representation which has been supported by consumer behaviour research conducted from various disciplinary perspectives (including economics, marketing and management science).

In more formal terms, the growth curve is commonly operationalised through a logistic function, giving rise to symmetrical S-shaped growth curve, which increases with the drivers of demand, from zero in the distant past to the saturation level of demand in the distant future.

An important illustration of this method within a transportation context is Tanner's (1978) pioneering research on car ownership forecasting¹⁰. To cite a specific example, in relation to the so-called *LR650 forecasts*, Tanner reports a saturation level of 0.45 for cars per head, along with the respective contributions of three factors in driving demand growth, namely increasing income (45%), reducing costs of motoring (15%), and 'other' factors changing over time (40%).

Relative to road, it is perhaps fair to say that rail has devoted rather less attention to the phenomenon of saturation, but this is perhaps reflective of empirical evidence concerning its prevalence in rail markets (see later discussion on this point).

The underlying reasons for saturation can be many, and due to factors on both demand and supply sides, but the present discussion will focus primarily upon the former. That said, and referring back to earlier discussion in section 3.1, it is worth noting that, through the inclusion of crowding as a quality variable, PDFH offers a simple representation of the influence of capacity on passenger rail demand.

When investigating the prevalence of demand saturation in empirical data, researchers have typically focussed upon the income elasticity of demand. In simple terms, demand saturation is considered to hold where the income elasticity of demand is zero, i.e. meaning that an increase in income will have no positive (or negative) impact on demand. More generally, researchers have explored the prevalence of a downward trend in the income elasticity, as might indicate the maturing of demand towards saturation.

The current PDFH v5.1 GDP elasticities are all conditional upon the constraints on car time and fuel cross elasticities, population elasticities and car ownership effects. For non-London flows, an innovation of PDFH v5.1 has been to distinguish between 'short' and 'long' term, with the former marginally higher than the latter.

This pattern of GDP elasticities reducing over time would support an argument that rail patronage is progressing towards saturation but, since they remain considerably in excess of zero, would also indicate that there remains scope for further market growth.

With regards to the different elasticities to/from London, the previous DfT-commissioned demand cap study by Bates *et al* (2013) commented: *'given incomes are higher in London than elsewhere, this could be indicative of a saturation effect. However, the greater focus of economic activity on London might actually have driven larger income elasticities of business trips to London. In terms of leisure trips, it is generally felt that increasing discretionary income will serve as a stronger incentive to short stay visits to London by rail than from London, not least because Londoners are more likely to use their cars for trips elsewhere whereas using a car for a trip to London is an unattractive proposition'*.¹¹

¹⁰ Tanner, J.C. (1978) Long-Term Forecasting of Vehicle Ownership and Road Traffic. *Journal of the Royal Statistical Society*. Series A (General), Vol. 141, No. 1 (1978), pp. 14-63

¹¹ Bates, J. *et al* (2013) Specifying the demand cap for rail. A report to the Department for Transport.

It is notable that, whilst distinguishing between short/long term and to/from London, PDFH v5.1 does not explicitly attribute these segmentations to saturation.

More generally, econometric studies of British passenger rail demand have consistently found that the market is continuing to grow. Notable among these studies is Arup and Oxera's (2012) comprehensive study, which concluded that: *'There is no evidence of market saturation. Assuming that the economy will grow according to trend in the long term, rail demand will continue to increase, and plans will have to be drawn up to cater for this growth'*.¹²

Slightly qualifying this conclusion, Bates *et al* (2013) commented: *'All in all, there appears to be little evidence to support the GDP elasticity falling as demand reaches saturation. However, it could well be that any analysis of how the GDP elasticity varies over time has faced a number of limitations that seriously impact on its ability to detect any demand saturation effects...'*

Such limitations could potentially include:

- un-modelled trend effects, such as cohort effects and structural change – although these phenomena would seem to be unavoidable hazards of econometric modelling;
- the failure to accurately isolate other external factors even when entered into models – this would seem to be a weakness of PDFH, given the failure to predict continued demand growth throughout the recent period of recession; and
- the use of regional GDP/GVA variables – this practice is advocated by PDFH, but the result has been considerably larger income elasticities for regional centres.

¹² Arup & Oxera (2012) Revisiting the elasticity based framework – update. Final report to the Department for Transport.

3 Historic Scheme Trends

This Chapter describes the results of our individual scheme analysis. The key here is to understand the profile of demand *before the 20 year mark*, and *after the 20 year mark*, to get an understanding of what happens to demand in reality after schemes are built. The analysis suggests that demand does not stop growing after 20 years, when the current demand cap is applied.

This Chapter has the following sub-sections.

- **Scheme Selection** – summarises how the schemes were selected.
- **Scheme Demand Profiles** – describes the demand profiles of the schemes, before and after 20 years.
- **Scheme Demand Growth vs Population Growth** – comments on the suitability of using forecast population growth as a predictor of demand growth.

3.1 Scheme Selection

Transport schemes were selected from those that were built 20 or more years ago, and for which time-series data was available (see **Table 2**). These are schemes available in our data bank and the degree to which they are representative of all transport schemes is unknown. Local major road schemes are underrepresented, for example.

Our analysis focuses on how schemes grow 20 years *after opening*, although the demand cap in rail tends to be applied 20 years *from the appraisal year*, rather than opening year, which may be five to ten years before the opening date itself. This is primarily due to being unable to review the business cases for all of these schemes within the time frame of this study, to understand when the cap year was applied. Appendix A shows the demand profiles for each of these schemes individually with commentary on significant events and drivers of change.

Table 2 – Schemes examined in summary analysis

Scheme	Opening Year	Mode
A19 Barlby Junction bypass (near York)	1987	Road
A19 Billingham Diversion (near Middlesbrough)	1983	Road
Dualling of A66 between Banks Gate and Bowes West	1993	Road
Dartford crossing (fully connected to M25)	1986	Road
Dartford crossing (west tunnel opens)	1963	Road
M11 Motorway (fully opened)	1980	Road
M23 between Hooley and Mertsam (J7-J8)	1974	Road
M25 between Micklefield to South Mimms (J19 to J23)	1986	Road
M25 Motorway (fully opened)	1986	Road
M3 between Sunbury to Lightwater (J1 to J3)	1974	Road
M40 between Stokenchurch to Waterstock (J5 to J8a)	1974	Road

M42 Southern Links (M5 to M42 J1)	1987	Road
M5 between Avonmouth Bridge (J18 to J19)	1974	Road
M65 between Nelson to Colne (J13 to J14)	1988	Road
M74 between Millbank and Nether Abington (J12-J13)	1991	Road
A616 Stocksbridge bypass (near Sheffield)	1988	Road
Docklands Light Railway (DLR)	1987	Rail
InterCity East Coast (Electrification)	1991	Rail
Glasgow Subway Re-opening	1980	Rail
Jubilee Line	1979	Rail
Manchester Metrolink	1992	Rail
Sheffield Supertran	1994	Rail
Tyne and Wear Metro	1980	Rail

Source: Arup

Some of the schemes examined have had significant capacity expansions or ridership-stimulating upgrades beyond the opening date. For example, the DLR has had numerous extensions over the years, motorway networks have been expanded, and feeder roads have changed in nature. The absence of truly unaltered schemes has made it more difficult to draw conclusions on the organic growth that schemes typically experience over the long term, however we have attempted to address this by removing a number of schemes from our sample in some of the analysis.

Some charts and tables in this section show values that ‘exclude schemes with significant further investment’, and these schemes are as described in **Table 3**.

Table 3 – Schemes with significant further investment

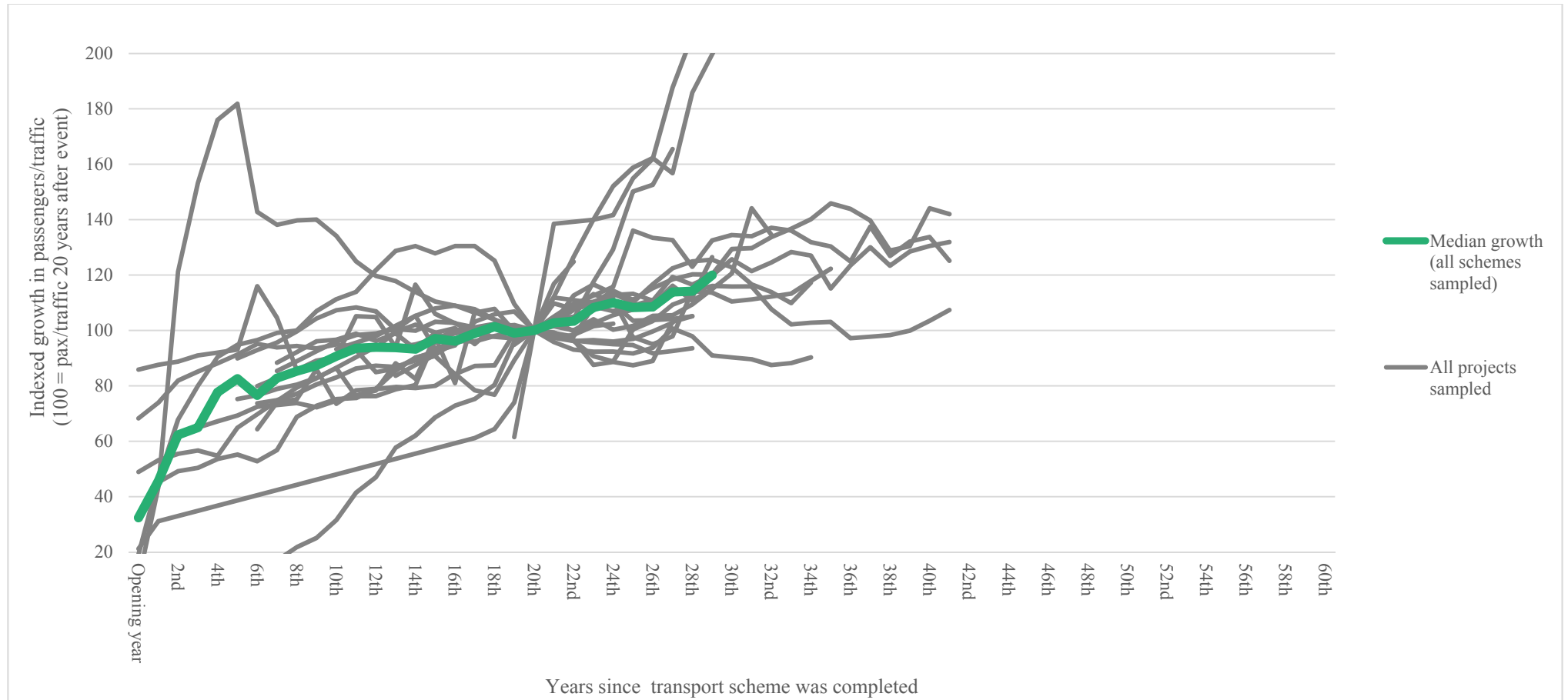
Scheme	Investment
Docklands Light Railway	DLR has been expanded numerous times since opening.
Jubilee Line	Jubilee Line was extended in 1999
Dartford Crossing	Since opening the Dartford Crossing has expanded to include a second tunnel and bridge crossing
M25 motorway	Since opening there has been a series of major junction improvements and widens
Manchester Metrolink	The Metrolink system has been expanded multiple times, most notably in 1999, 2002, 2014.

Source: Arup

3.2 Scheme Demand Profiles

The key point here is that demand for *transport schemes tends to continue to grow 20 years after opening*, as shown in **Figure 16**. It shows in green the median growth profile of all schemes sampled across road and rail and it demonstrates that a typical scheme tends to continue growing between their 20th and 35th year. Beyond the 35th year we have insufficient data to show clear result.

Figure 16 – Median growth profile for all schemes sampled



Source: DfT & Rail Industry Monitor

Our analysis indicated that road and rail projects had similar growth rates *once schemes that have had further significant investment were excluded*. The fact that road and rail schemes have similar ‘organic’ growth profiles after 20 years could indicate that a consistent approach to growing demand in later years could be acceptable for both rail and road. **Table 4** illustrates these findings with the mean outcomes.

Table 4 – Mean annual demand growth rates by age of project and mode

	Y4-Y19	Y20-Y35	Y36-Y60 ¹³
<i>All projects sampled</i>			
Road	2.5%	1.9%	0.5%
Rail	3.8%	6.1%	Insufficient data
<i>All projects sampled – excluding schemes with significant further investment¹⁴</i>			
Road	2.8%	1.3%	0.6%
Rail	1.1%	1.0%	Insufficient data

Source: DfT & RIM

Note that when excluding schemes with significant further investment, we are excluding schemes which have been in high growth places.

3.3 Scheme Demand Growth vs Population Growth

We consider here whether a simple rule could be utilised as a way of forecasting long term demand growth. The rule considered is that ‘forecast population growth can be used on its own as an approximation for demand growth’.

- **Rationale.** Using population-based growth is another approximation to considering long term demand growth, like the current demand cap guidance. Using population growth as a tool for demand forecasting in the post-modelling period (Year 20) should only be applied if there is certainty or evidence that this alone is what drives long term growth in demand.
- **Evidence.** However, no evidence exists to suggest that population growth can be applied as a general rule to all schemes to forecast long term demand growth. Population growth may be an appropriate proxy for demand forecasting for certain schemes, but not for others. We know, based on our evidence from previous road and rail schemes, that for some schemes, the rate of demand growth has far outstretched the comparable rate of population growth over the same period.

The population growth forecasts also provide a limited indication as to the distribution and intensity of the growth. Our view is that an overall general approach like growing demand based on population growth isn’t appropriate.

¹³ We have limited data for the profile of demand for Y36-Y60.

¹⁴ Schemes excluded include: DLR, Jubilee Line, Dartford Crossing, M25 motorway and Manchester Metro Link

4 Simulation Model

This chapter describes our approach to simulation modelling and includes the following sub-sections.

- **Simulation Model** – an overview of the simulation models built for road and rail and results of the sensitivity tests.
- **Project Case Study** – briefly describes a real project example which is discussed further in Appendix B.

Table 5 summarises the tests that were undertaken using the simulation model.

Table 5 – Tests considered

Test Name	Mode	Rationale
Highway Demand Cap	Road	Represents the impact on the PVB (£) of capping at different points of the road scheme lifecycle
Highway Generalised Cost Cap	Road	Developing a more accurate assessment of where the demand cap should be placed by relating the cap to supply/demand (GC)
Highway Uncertainty Equivalent	Road	Assess the impact of applying a certainty equivalence to the benefits stream to account for issues regarding uncertainty.
Rail Demand Cap	Rail	Represents the impact on the PVB (£) of capping at different points of the rail scheme lifecycle
Rail Generalised Cost Cap	Rail	Developing a more accurate assessment of where the demand cap should be placed by relating the cap to supply/demand (GC)
Rail Uncertainty Equivalent	Rail	Assess the impact of applying a certainty equivalence to the benefits stream to account for issues regarding uncertainty.
Road Case Study (Appendix B)	Road	To test the impact of the demand cap removal on a real-world project

Source: Arup

4.1 Simulation Model

A simulation model has been developed to illustrate the change in benefits over time based on the supply-demand principles discussed previously in Section 1.6. This model calculates the change in the demand curve and the supply curve for a particular road or rail scheme over time. By calculating these changes, the model then determines the discounted benefits for the scheme using the ‘Rule of Half’ principle and the discounting values provided in WebTAG.

Separate models were developed for the road and rail contexts. It is important to note that both the road and rail simulations are illustrative, and focus on the user benefit associated with transport improvements using the ‘Rule of Half’ measure. In these simulation models, the schemes are considered in isolation, with no consideration of alternative routes or modes; whereas in a typical transport demand model, schemes are considered within the context of the wider network.

No account is taken of indirect tax effects of induced traffic or wider economic impacts which would be considered in a real-life appraisal. There is also no consideration of environmental or safety impacts.

We have made use of the simulation models to illustrate the effects of allowing demand to grow as determined by long term projections of the main drivers of demand, but it is also constrained to reflect supply side capacity limitations. In other words, we have taken a simplified version of a conventional transport model to calculate the benefits of a hypothetical scheme. The model has been set up to ensure that it continues to converge to an equilibrium in both the do-minimum (DM) and do-something (DS) scenarios over the 60 year appraisal period. There is nothing internal to the demand modelling process which is equivalent to a minimum speed or maximum flow cut-off. Alternative approaches to capping demand were then introduced into the model and the impact of each option on benefits is described. Note that in practice, the working of a capping process would need to be considered in conjunction with the working of any internal modelling rules.

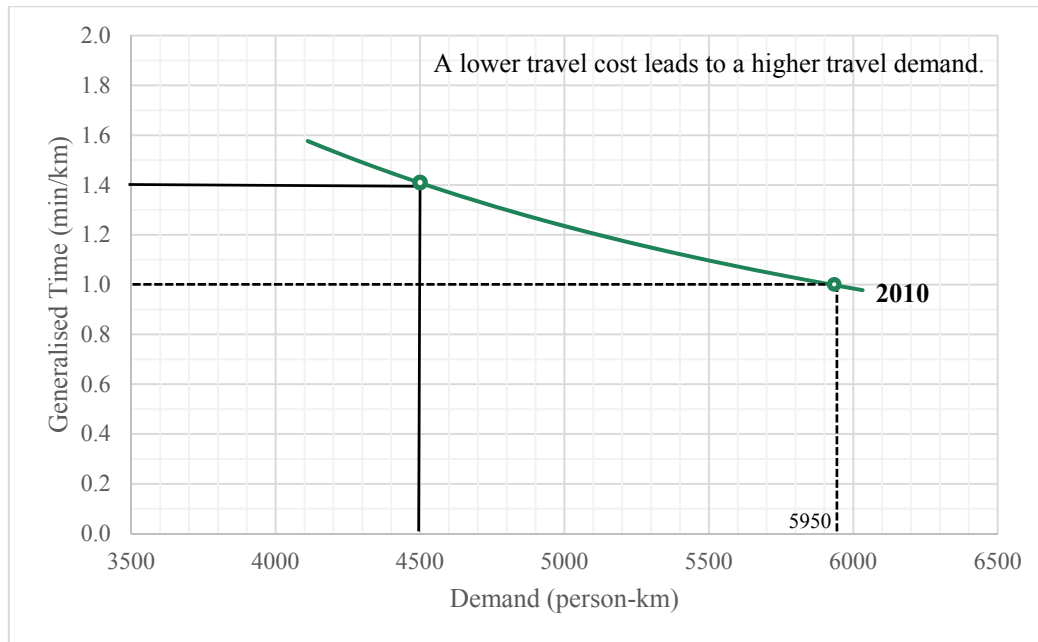
Note that a limitation of the simulation model is that it may not be fully representative of all the phenomenon modelled in complex transport models used in appraisal. However it was considered by the study team as being fit for purpose for this work – i.e. that it would capture many material aspects of benefits quantification. It's also a strength of the simulation model that it is relatively easy to test many scenarios which would be costly to do with a more complex modelling suite.

In the following sections, we have restricted our analysis to a single hypothetical scheme for each mode to demonstrate the effects of options for capping on demand and benefits.

4.1.1 Road Simulation Model

This section describes the components of the demand and supply curves in the model, the assumptions of the road simulation, and the interpretation of the simulation results. This model calculates the demand and supply curve for each year of the appraisal period.

The demand curve represents what the peak-hour demand for the highway will be in a given year depending on the cost to use the highway. The demand curve can be thought of as the 'user response' to changes in cost. This cost, labelled as 'generalised time', is calculated as the sum of travel time plus the fuel cost for the vehicle converted to time units using a value of time. **Figure 17** illustrates an example demand curve for the base year 2010.

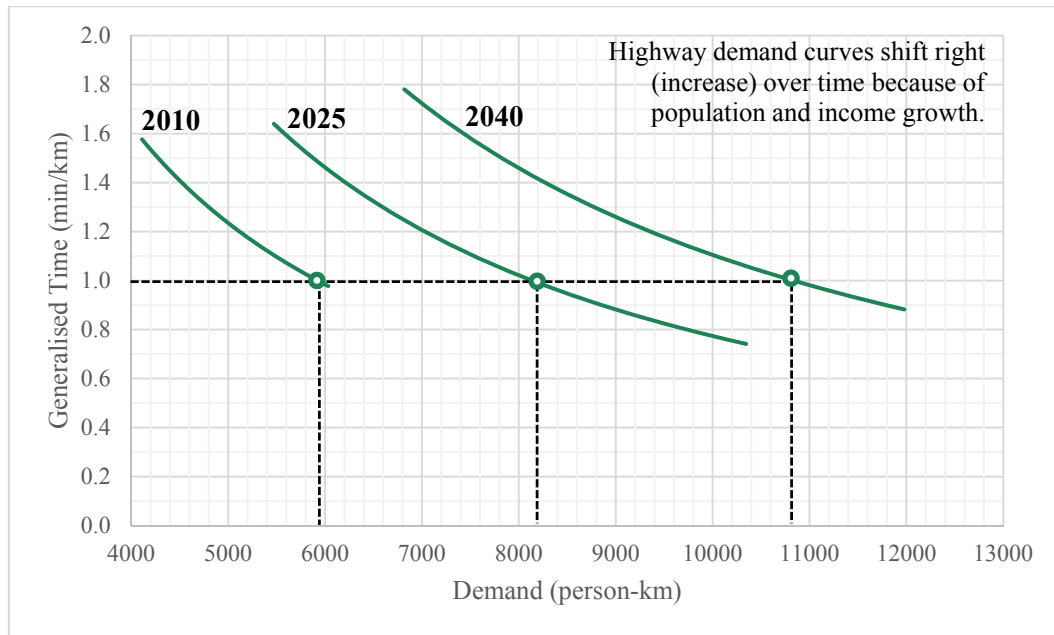
Figure 17 – 2010 Base Year Demand

Source: Arup

This demand curve illustrates the highway demand over a range of cost (generalised time) in the base year 2010. As an example, if the cost to travel the highway in the base year is 1.4 min/km, then the demand for the highway is 4500 person-km. If the cost to travel the highway is lowered to 1.0 min/km, then the demand will increase to 5950 person-km. The slope of the curve reflects of how sensitive the demand volume is to changes in generalised time, all else constant. The chosen generalised time elasticity is compatible with the recommended WebTAG fuel cost elasticity of -0.3.

Over the course of a 60 year appraisal period, the demand curve will not remain static, but it will shift with changes in external factors, which in the road model, include population and income. Using WebTAG inputs for population and value of time (as a proxy for income), these factors are expected to increase over a 60 year period. Therefore, the demand for the highway is also expected to increase; all else does not stay constant over time. The simulation model recalculates the demand curve for each year of appraisal period taking into consideration the increases in population and value of time. We have assumed an income elasticity of 1, which may be too high: this should be borne in mind in assessing the results. **Figure 18** shows three of these demand curves, one for 2010, 2025 and 2040.

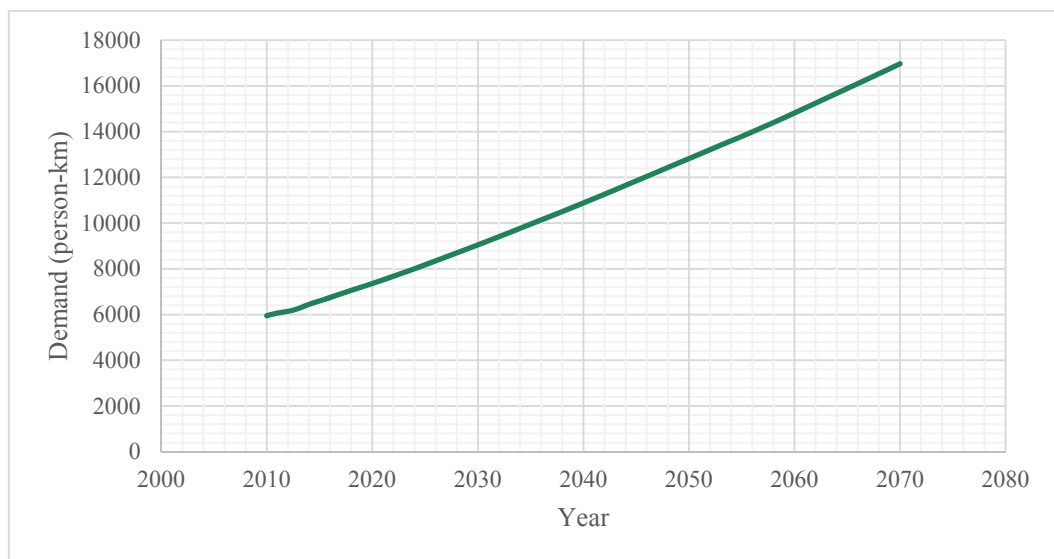
Figure 18 – 2010, 2025 and 2040 Highway Demand Curves



Source: Arup

As shown in the above figure, the demand curve shifts towards the right over time due to increases in population and income. These underlying external factors will cause the demand to increase substantially. If the cost to travel the highway is 1.0 min/km, then while the demand in 2010 is 5,950 person-km, by 2025 the demand would be 8,200 person-km, and by 2040 the demand would be 10,800 person-km. Assuming that there is no change in the generalised time, which means that both the cost and the time to travel on the highway remains fixed, we can show the overall growth in the demand based on the changes in population and income in **Figure 19**. This may be thought of as latent demand growth.

Figure 19 – Highway Demand for a Fixed Cost (1.0 min/km)

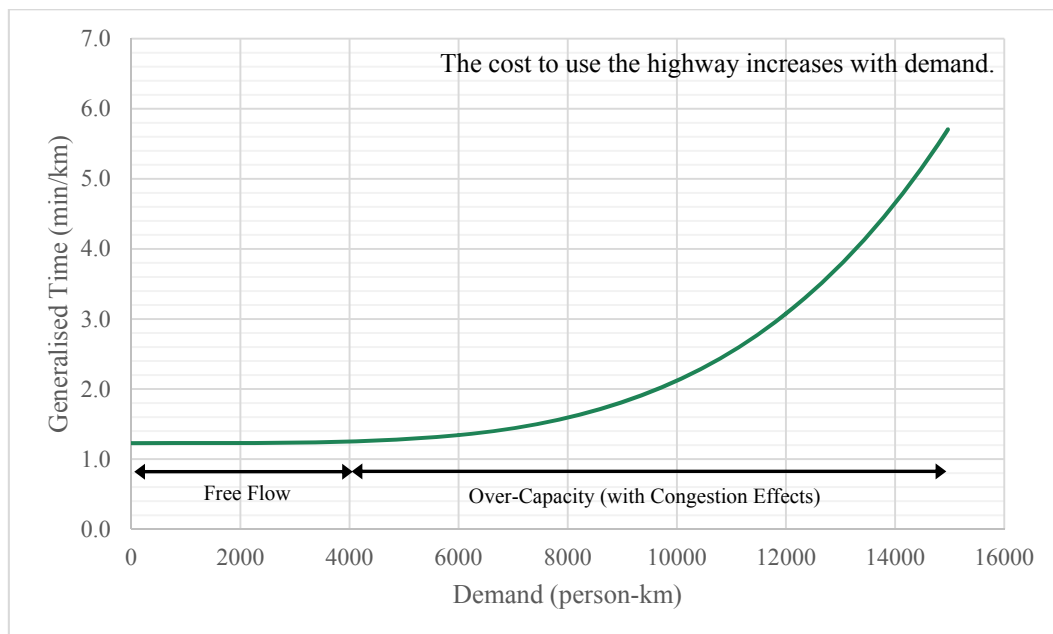


Source: Arup

However, calculating the demand curve is only half of the story. Up until this point, we have considered only demand growth, assuming that the capacity is able to handle all growth. However, we must also consider how capacity affects demand. To do this, the model must also include a supply curve.

The supply curve in the road model can be considered as the ‘infrastructure response’ to changes in demand, as it illustrates the generalised cost to use the highway over a range of demand. Typically, as the demand for a highway grows, the time it takes to travel the highway should also grow as the highway becomes more congested. The supply curve in the road model is calculated using a volume-capacity relationship known as the Bureau of Public Roads (BPR) Function, which is typically used in traffic assignment models. This function suggests that traffic should travel at the ‘free-flow speed’ if the demand for the highway is much less than the capacity of the roadway. As the highway demand approaches and exceeds the highway capacity, congestion effects take place, causing traffic to travel at much slower speeds, which means travel times are significantly higher. **Figure 20** illustrates a supply curve in terms of generalised time, which combines both travel time and the fuel cost converted into time units.

Figure 20 – 2010 Highway Supply Curve



Source: Arup

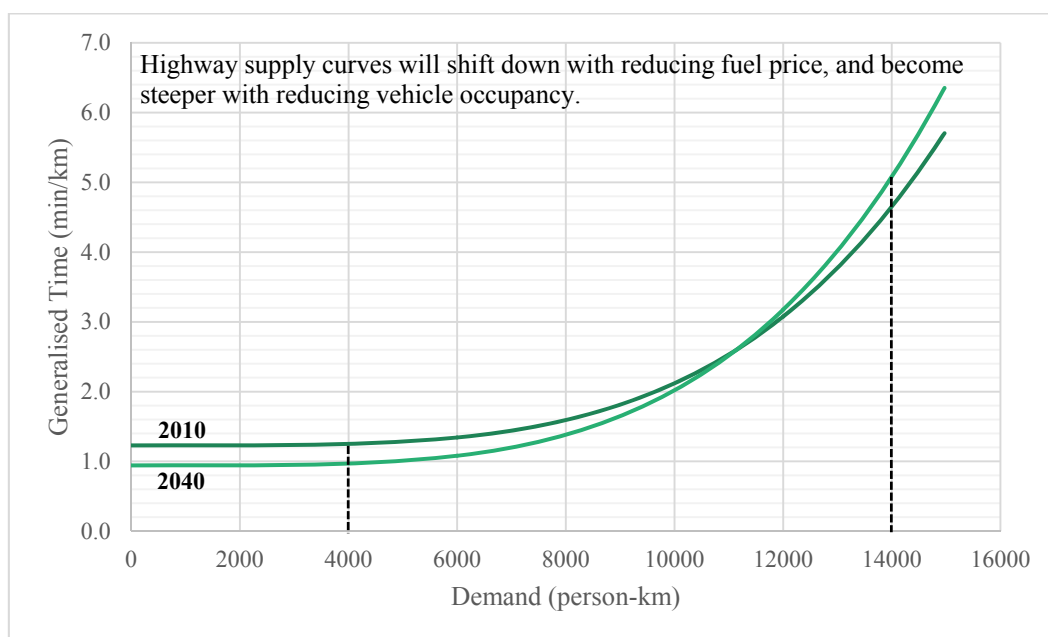
The above supply curve represents the combined cost of time and fuel to travel on a 1 kilometre highway with a free-flow speed of 80 kph (50 mph) and a capacity of 4000 person-km given the fuel and time costs pertaining in that year. The above curve is derived using 2010 values for those variables. The far left of the chart is the highway in a free-flow state. Even if the demand is very low (< 2000 person-km), the total cost to use this highway is 1.2 min/km. The highway remains at this state until approximately 4000 person-km, the assumed capacity of the highway, where the cost starts to rise substantially with demand. Note that fuel prices and values of time stay constant in one given year, so any increases in

cost for an individual supply curve is purely based on higher travel times due to congestion factors.

Similar to the demand curve, the supply curve does not remain static over time but can change due to external factors. These factors include WebTAG inputs for fuel cost, car efficiency, and value of time forecasts from 2010 to 2070. Over these years, the forecasts suggest that fuel prices will decrease, car efficiencies will improve, and values of time will increase. The combination of these effects means that the base cost (excluding travel time) to travel on the highway should decrease over this period. Therefore, future supply curves should be located lower on the graph than the base year supply curve for smaller values of demand, and then to the left of the base year supply curve for higher values for demand.

Changes in car occupancy also affects the shape of the supply curve. Based on WebTAG inputs, the number of passengers in a vehicle is expected to decrease over the period between 2010 and 2070. A lower car occupancy suggests that, as more vehicles try to use the highway, the highway will react more quickly to increasing demand (measured throughout this analysis in terms of person kilometres). Therefore, future supply curves should become steeper more quickly than the base year supply curve, and congestion effects are experienced more quickly than in the base year. These effects are shown in **Figure 21**, which illustrate two supply curves: one for 2010 and another for 2040.

Figure 21 – Highway Supply Curves for 2010 and 2040



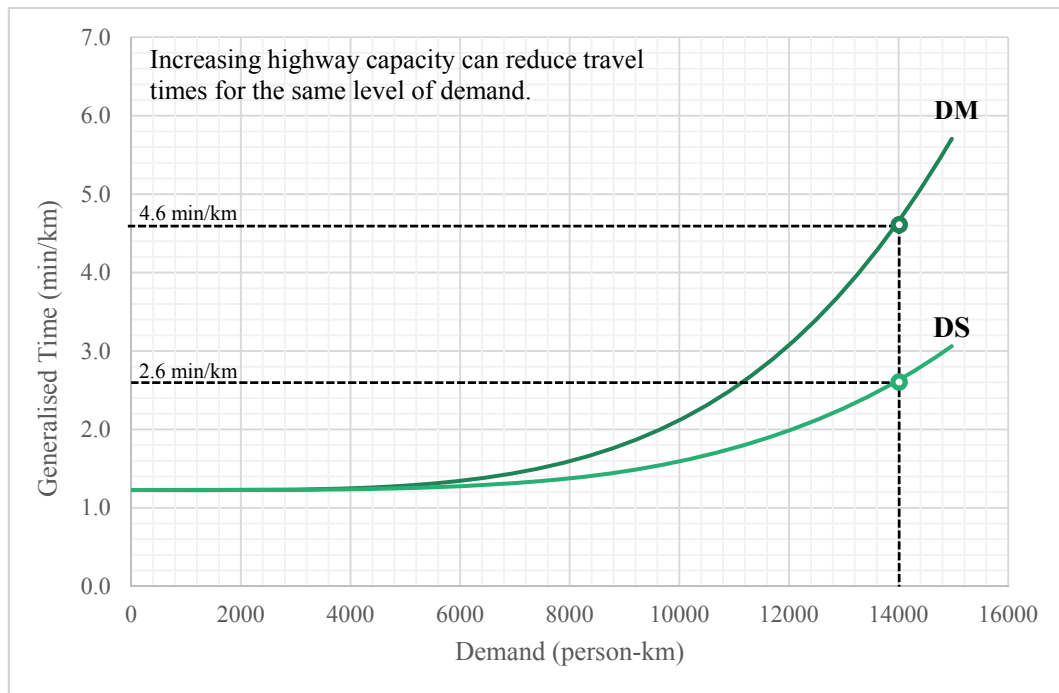
Source: Arup

The above figure shows the combined effects of changes in fuel cost, and changes in vehicle occupancy. For an uncongested road (i.e. in this case demand less than 4000 person-km), reductions in fuel cost explains the overall reduction in cost (e.g. 1.2 min/km in 2010 to 1.0 min/km 2040). In contrast, as the road becomes congested at higher levels of demand, the changes in vehicle occupancy explains why the 2040 cost is higher than the 2010 cost. For the same demand (e.g. 14,000

person-km), reductions in vehicle occupancy changes the overall cost to travel the highway from 4.6min/km in 2010 to 5.0 min/km in 2040.

The supply curve can also change based on the features of the proposed transport scheme. Suppose that this 1 kilometre highway, which has a current capacity of 4000 person-km, has a potential scheme to increase its capacity through a road widening and the addition of a lane. With this scheme, the capacity increases by 25% to 5000 person-km. **Figure 22** shows the 2010 supply curves for the existing condition, known as the Do Minimum (DM) supply curve, and the proposed scheme known as the Do Something (DS) supply curve.

Figure 22 – 2010 Highway DM and DS Supply Curves



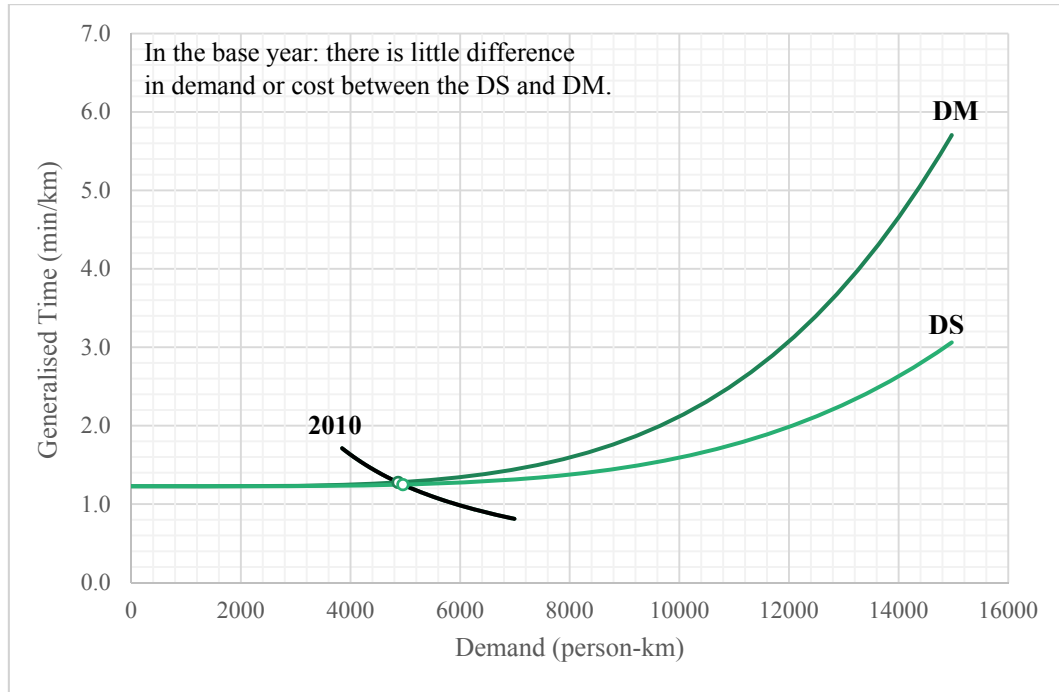
Source: Arup

In the above figure, both DM and DS curves have the same ‘free flow’ costs of 1.2 min/km for demand that is less than 4000 person-km. This is expected as these curves are representative of the same year. However, at higher levels of demand, the DS curve has a lower cost than the DM curve. As an example, at 14,000 person-km, the existing highway (DM) has a cost of 4.6 min/km, while the expanded highway (DS) has a cost of 2.6 min/km. This result is expected as the DS highway has a 25% higher capacity than the DM highway, which means that the DS highway is able to handle much more demand until it reaches the same level of congestion. As mentioned previously in this section, the supply curves will shift in each year in response to external factors.

The model now contains three sets of curves: one for Demand, one for DM Supply, and one for DS Supply. Within each set, there are up to 60 curves: one for each year of the appraisal period. As discussed above, the demand curve can be thought of the ‘user response’ as changes in cost influences the overall demand, while the supply curve can be thought of as the ‘infrastructure response’ as changes in demand affects the cost to use the highway.

The focus now shifts to how these curves interact. The demand and the two supply curves have been plotted with demand on the x-axis, and cost (generalised time) on the y-axis. Therefore, all three curves can be plotted on the same graph as shown, for the base year 2010, in **Figure 23**.

Figure 23 – 2010 Highway Demand, DM Supply and DS Supply Curves

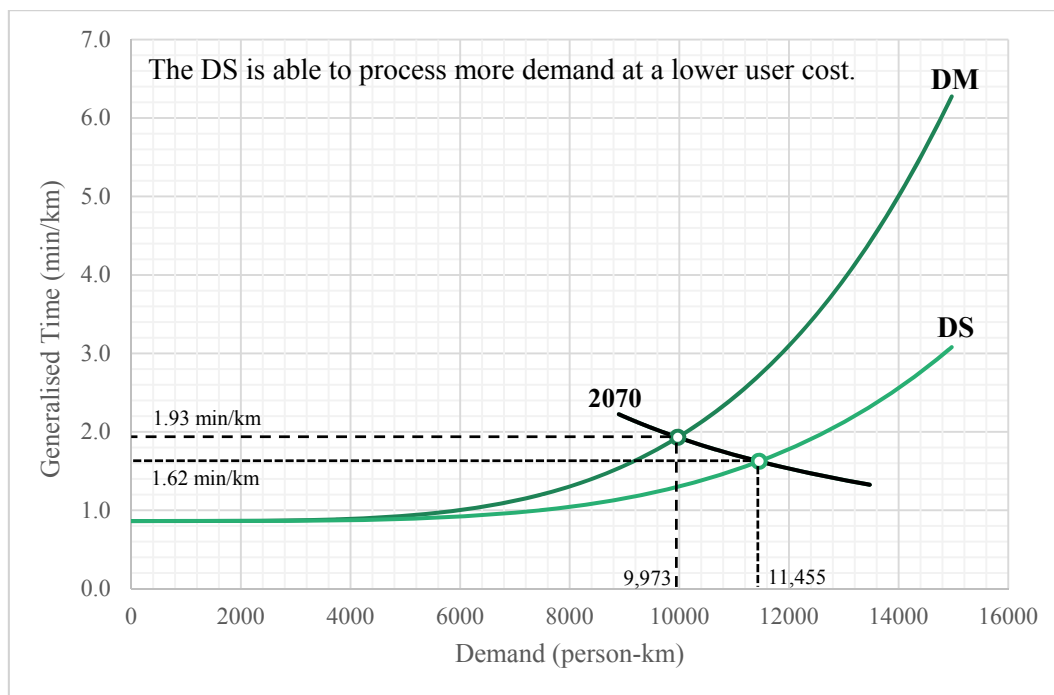


Source: Arup

The point on the graph at which the demand curve and the supply curve intersect is known as the demand – supply equilibrium point, and it is where the demand and cost is expected to be for the given year. In **Figure 23**, the demand curve intersects the DM Supply curve at a demand of 4,866 person-km, and a generalised time of 1.28 min/km, which corresponds to the assumed existing operating speed of 75 kph (45 mph). This can be interpreted as the base year demand and cost for the existing highway without any additional capacity. Note that the intersection point of the demand curve and the DS Supply curve is in a similar location. This result suggests that the proposed highway expansion to provide an additional 25% capacity has very little travel cost savings in the base year, as the existing demand is close to the free flow capacity.

The model repeats these calculations for each year of the appraisal period as each curve is expected to shift based on external factors (for demand: population and income, and for supply: fuel cost, vehicle efficiency, and vehicle occupancy). By the 60th year of the project, the demand DM Supply and DS supply curves will have shifted as illustrated in **Figure 24**.

Figure 24 – 2070 Highway Demand, DM Supply and DS Supply Curves



Source: Arup

By 2070, population and income growth will cause the demand curve to shift right, indicative of an increase in demand. As well, improved fuel efficiencies and reduced fuel cost will shift the supply curves down, representing a lower base cost to use the highway. The combination of these changes cause the demand curve to intersect the DM Supply curve and the DS Supply curve at different points, where there is a greater difference in demand and cost, as summarised in **Table 6**.

Table 6 – 2070 DM and DS Equilibrium Demand and Generalised Time

Scenario	Demand (person-km)	Generalised Time (min/km)
Do Minimum (DM)	9,973	1.93
Do Something (DS)	11,455	1.62

Source: Arup

This result suggests that if no changes are made to the existing highway, then congestion effects experienced from high demand will significantly increase travel times. If the proposed scheme to expand capacity by 25% is undertaken, then the highway would be able to process more demand at a lower cost.

In each year, the transport benefits (in terms of cost savings) that can be realised by implementing the proposed highway scheme is calculated using the difference of the intersection point of the Demand Curve and the DM Supply Curve, and the intersection point of the Demand Curve and the DS Supply Curve. This is also known as the ‘Change in Consumer Surplus’, which is calculated using the Rule of Half, which was discussed in Section 1.6 of this report, as well as in WebTAG Unit A1.3, Section 2.

As an example, the process to calculate the change in consumer surplus for 2070 is outlined below using a simplified version of Equation (2) (refer to Section 1.6.2 for the complete version of the equation), and using the values for the variables summarised in **Table 7**.

$$S = \frac{1}{2} \times (T' + T) \times (C - C') \quad (7)$$

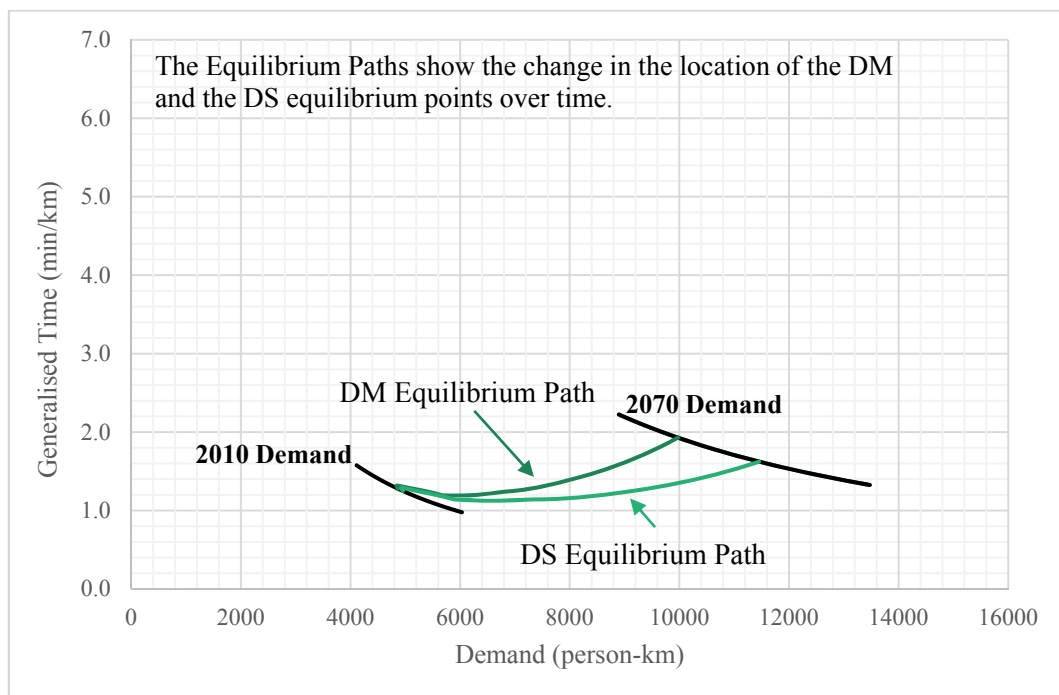
Table 7 – Values for Example Rule of Half Calculation

Variable	Description	Value
T'	DS Demand	11,455 person-km
T	DM Demand	9,973 person-km
C'	DS Cost	1.62 min/km
C	DM Cost	1.93 min/km

Replacing the variables with the values in the above table will lead to a calculated change in consumer surplus value of 3321 person-minutes. This can be converted to a monetary value using a value of time.

When appraising the potential scheme over 60 years, the transport benefits depend on the difference between the Do Minimum and the Do Something equilibrium points. However, the location of these equilibrium points will shift each year as the factors that generate the demand curve and the supply change. Therefore, the equilibrium points and the benefits must be calculated for each of the 60 appraisal years, taking into account the change in the demand and supply curve factors in each year. **Figure 25** shows the path of the equilibrium points for both the Do Minimum and the Do Something scenarios as it changes over 60 years.

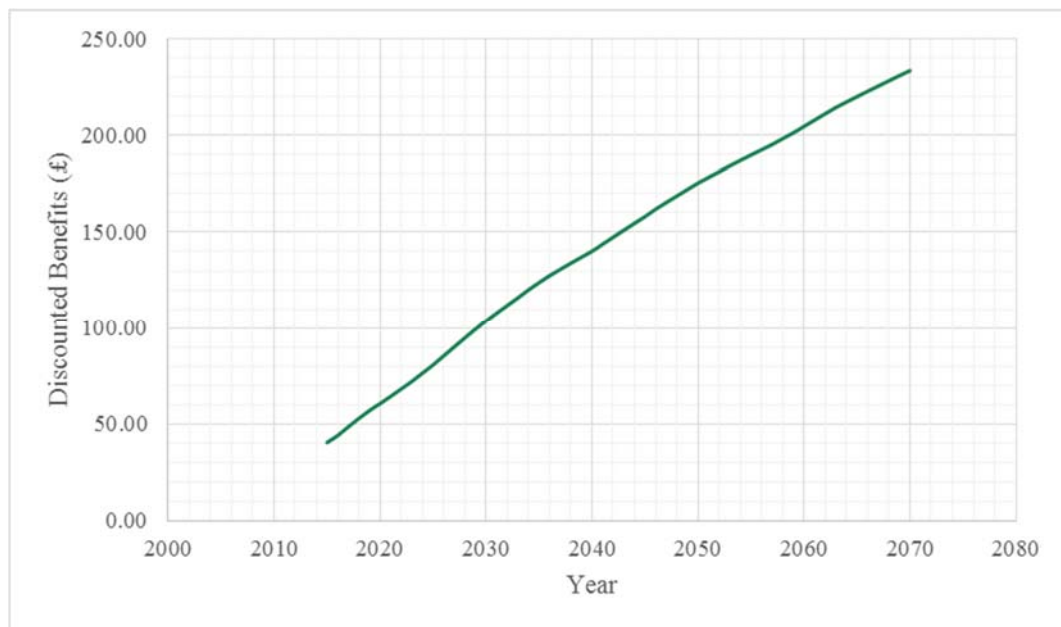
Figure 25 – Highway Do Minimum and Do Something Equilibrium Paths



Source: Arup

The difference between the two scenarios is very small in the base year because the operating speed of the Do Minimum is only marginally below the free flow speed, but by 2070, the generalised time savings from the Do Minimum to the Do Something scenario is much higher. In appraisal cases where congestion is a significant feature in the base year, demand and the equilibrium path would start out further to the right on the diagram. In this particular example, discounted benefits per annum rise throughout the entire sixty year appraisal period. Using the Rule of Half and the discount rates provided in WebTAG, the discounted benefits for the scheme are calculated and presented in **Figure 26**. Assuming that the scheme opens in 2015, the total present value of benefits [PVB] for the scheme is £8,111. It is noteworthy that year on year benefits continue to increase, albeit at a slowly decreasing rate, up to the 60th and final year of the appraisal. Such a trajectory of benefits might be interpreted as representative of a significant residual value attributable to the scheme at the end of the appraisal period.

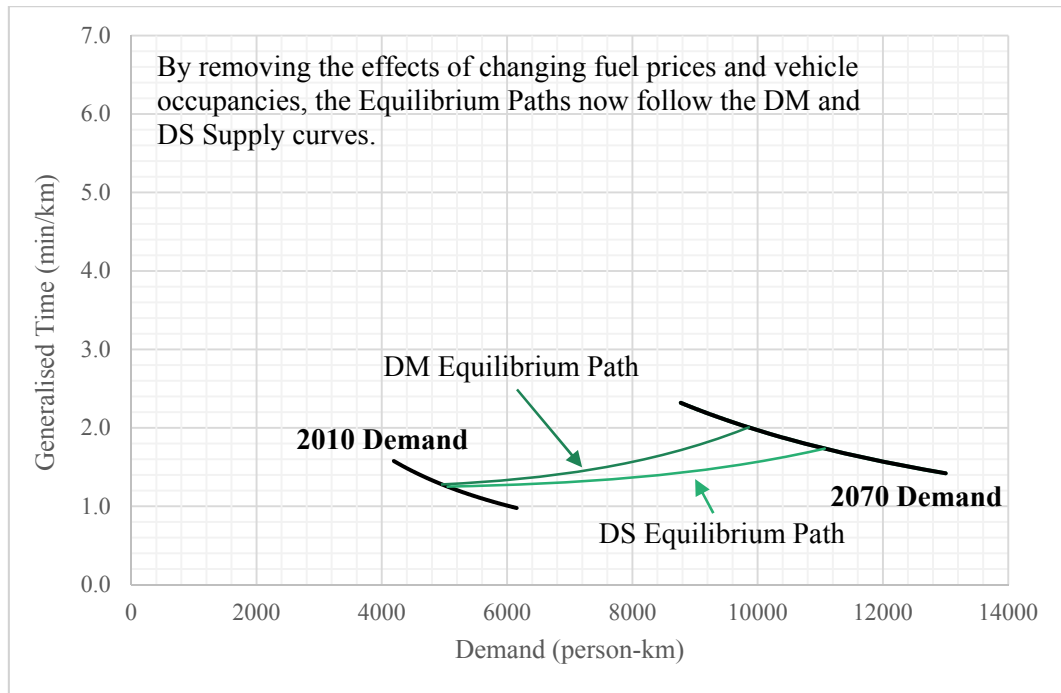
Figure 26 – Example Highway Discounted Benefits over the Appraisal Period



Source: Arup

The simulation model is then used to test the effect of capping demand on the discounted benefits. In order to avoid some of the complexities associated with the exogenous factors affecting the supply curve, and to provide better comparability with the rail simulation model, we have ignored the WebTAG changes in occupancy and fuel prices when testing the effect of different options for a cap. As a result the DM and DS Supply curves are assumed to be fixed in time, and the equilibria move along with them in line with shifts in the demand curve, as in **Figure 27**.

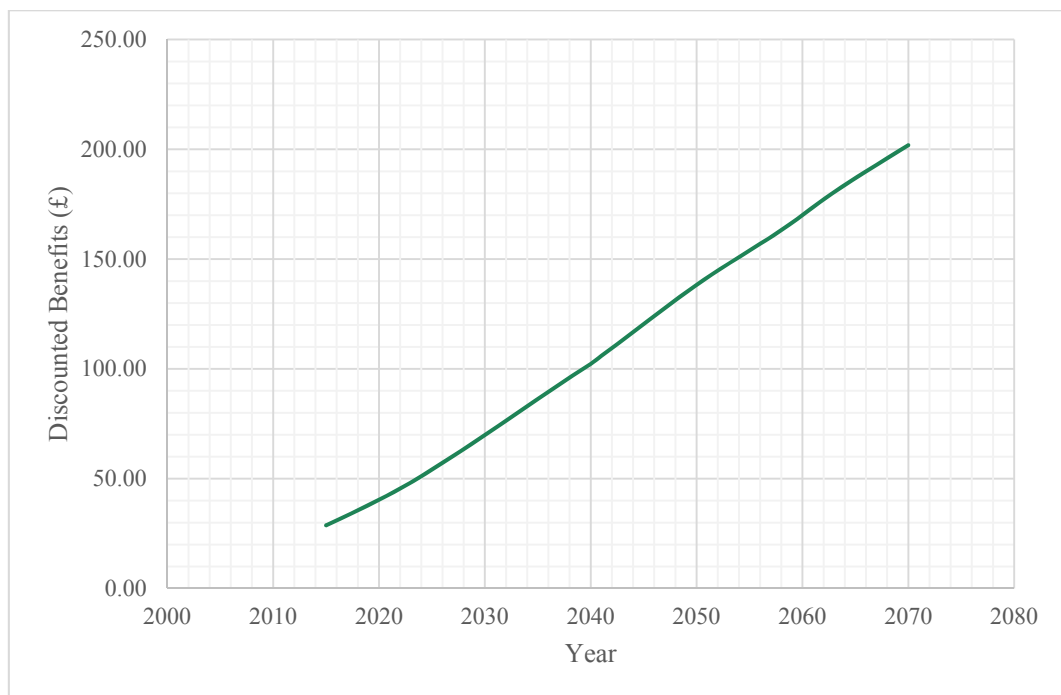
Figure 27 – Equilibrium Paths for Uncapped Highway Demand



Source: Arup

With this new assumption, the uncapped highway benefits have been recalculated. The discounted benefits over the 60 year period is summarised in **Figure 28**.

Figure 28 – Discounted Benefits for Uncapped Highway Demand



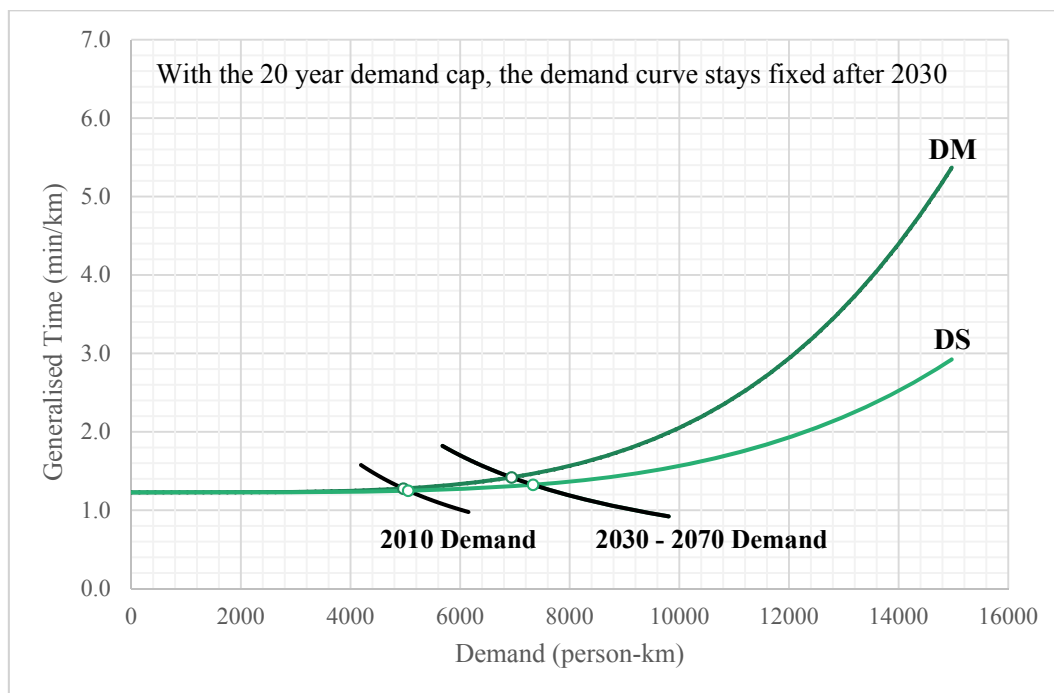
Source: Arup

Note that the PVB for this uncapped scenario is £6,303, which is lower than the previously discussed example, but this is expected as there are no benefits derived

from reductions in fuel cost and improvements in fuel efficiencies. This provides an appropriate baseline for comparison to the capped scenarios.

The first type of cap tested is a year cap, where **all** exogenous demand growth (based on changes in population and income) is capped after the x^{th} year of appraisal. After the demand cap is introduced, the demand curve is frozen, and does not shift further to the right. An example of the demand-supply curves for a 20 year cap scenario is illustrated in **Figure 29**.

Figure 29 – Highway Demand and Supply Curves for a 20 year cap



Source: Arup

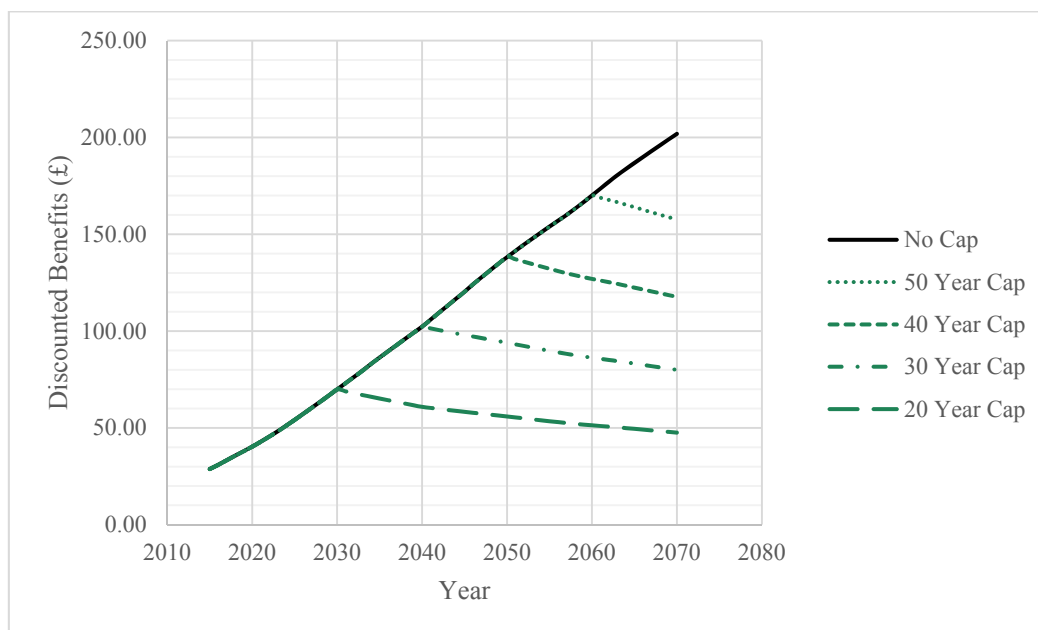
The number of years that were tested include 20 years, which is consistent with WebTAG guidance, as well as 30, 40 and 50 years after appraisal. All other parameters remain the same for this highway scheme. In each capped scenario, the discounted benefits would grow at the same rate as the ‘No Cap’ line up until the introduction of the demand cap. The benefits realised after the cap would be significantly less, as summarised in **Table 8**. The change in the discounted benefits curve is shown in **Figure 30**.

Table 8 – Highway PVB for Demand Cap with Years-based Approach

Scenario	Uncapped Period PVB (£)	Capped Period PVB (£)	Total PVB (£)
No Cap	£ 6,303	--	£6,303
50 Year Cap	£4,252	£1,804	£6,056
40 Year Cap	£2,728	£2,676	£5,404
30 Year Cap	£1,543	£2,802	£4,345
20 Year Cap	£697	£2,325	£3,022

Source: Arup

Figure 30 – Highway Discounted Benefits for Years-based Demand Cap



Source: Arup

Note that the cap is implemented by freezing the position of the demand curve. This is consistent with the approach taken in TUBA, where the output demand (and cost) is fixed. In both cases, the benefits increase according to the assumed growth in value of time. Since this growth is less than the discount rate, the discounted benefits fall subsequent to capping.

A variation of this test considers a year cap where demand growth as determined by income growth is capped, but it still responds with an elasticity of unity to changes in population. The resulting discounted benefits graph, shown in **Figure 31**, has higher benefits than realised under the scenarios in which all exogenous demand is capped as the demand will increase with population growth. The resulting present value benefits will also be higher, as summarised in **Table 9**.

Table 9 – Highway PVB for Income Only Years-based Demand Cap

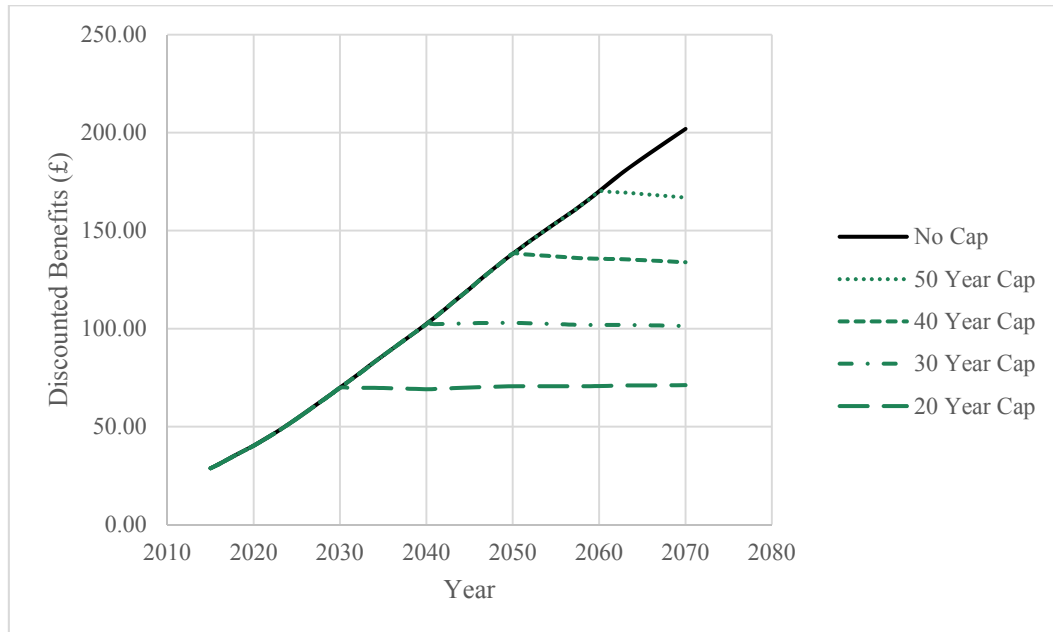
Scenario	Uncapped Period PVB (£)	Capped Period PVB (£)	Total PVB (£)
No Cap	£6,303	--	£6,303
50 Year Cap	£4,252	£1,855	£6,107
40 Year Cap	£2,728	£2,854	£5,582
30 Year Cap	£1,543	£3,171	£4,714
20 Year Cap	£697	£2,883	£3,580

Source: Arup

A feature of a years-based demand cap is the significance of the benefits in the final appraisal year. A well behaved benefits function might be expected to fall to a low value by the final year of the appraisal, justifying the choice of the appraisal period at the end of which the benefits might be expected to be small. Yet, as

Figure 30 and **Figure 31** show, the function reaches a cliff edge in the final year, with benefits significantly above those in the years immediately after opening.

Figure 31 – Highway Discounted Benefits for Income Only Years-based Demand Cap



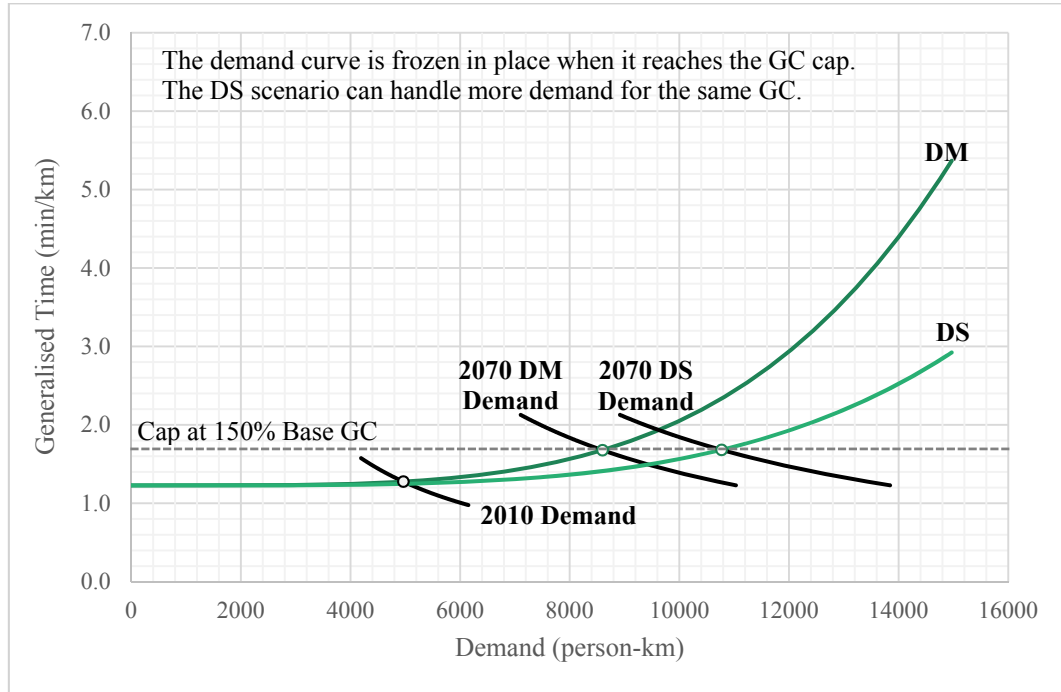
Source: Arup

A second test explores a variant of the demand cap that relies on the interaction of the demand and supply curves for the scheme. This variant is an attempt to be more accurate on when the demand cap should take place (i.e. the benefit of the Do Something scenario over the Do Minimum scenario should decrease to zero as the two scenarios become similar and the two have the same user outcomes). The cap that was tested was a limit based on the future generalised cost to travel on the scheme. Conceptually, we can argue that it might not be safe for equilibrium generalised cost to exceed base year generalised cost (GC) a threshold of x per cent, since beyond that zone, too many uncertainties and credibility issues arise. Above this threshold of base year generalised cost, which is indicative of heavily congested conditions, it might not be reasonable to model higher demand as travellers might consider other alternatives (other than the scheme being tested). Therefore, we cut off demand according to some threshold rule defined by the percentage of base year DM GC.

For the highway scenario, we have interpreted the generalised cost for capping as referring to the time units after removing fuel costs. This means that if the travel time on the highway exceeds a certain threshold of the base travel time, then the demand factors would be capped. Note that this will normally mean that the Do Minimum demand is capped prior to the Do Something demand. Once the Do Minimum demand is capped, the Do Minimum demand curve does not shift any further to the right. However, the additional capacity of the Do Something scenario allows more vehicles to use the highway before the generalised cost hits the threshold. The equilibrium for the Do Something scenario will shift upwards and to the right until the demand is capped. At this point, the Do Something

generalised time will be the same as the Do Minimum generalised cost. The equilibrium paths are illustrated for an example capping scenario (capped at 150% of base GC) in **Figure 32**.

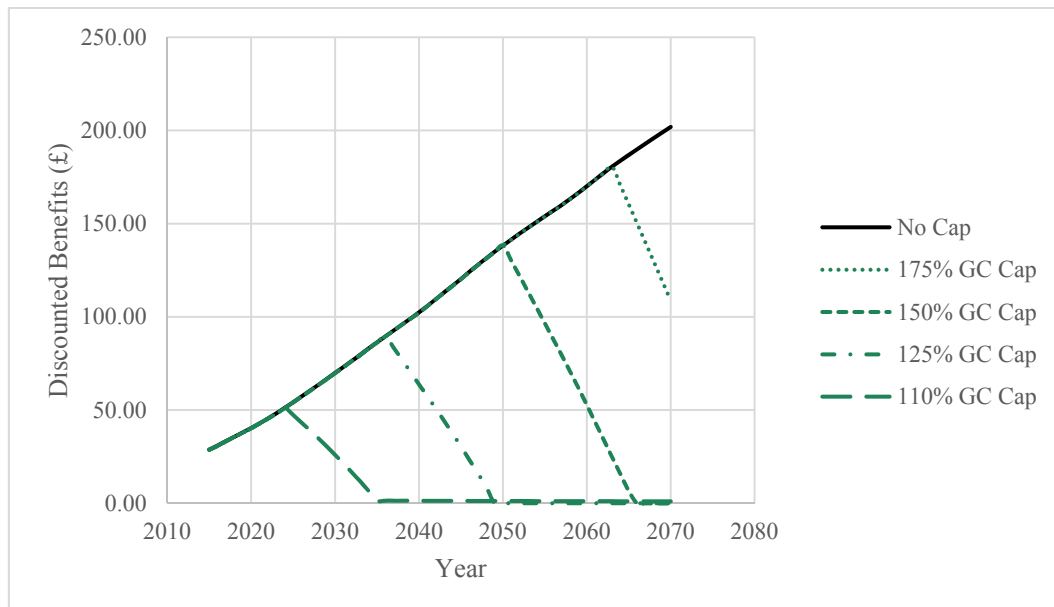
Figure 32 – Highway Demand and Supply Curves for a 150% GC Demand Cap



Source: Arup

Under this variant, a demand cap was introduced for all exogenous growth when the generalised cost exceeded a threshold of the base year generalised cost. For this test, the threshold was 110%, 125%, 150%, and 175% of the base year generalised cost. The resulting discounted benefits graph is in **Figure 33**.

Figure 33 – Highway Discounted Benefits for Generalised Cost-based Demand Cap



Source: Arup

As illustrated in **Figure 33**, there are three general phases that take place with the generalised cost cap.

- The first phase is the uncapped phase, where both Do Minimum and Do Something generalised cost are below the threshold. The benefits build up as the difference between the two equilibrium points increases.
- The second phase is when the generalised cost for the Do Minimum scenario has reached the threshold, while the generalised cost for the Do Something scenario is below the threshold. During this phase, the difference between the two equilibrium points decreases, and the travel time savings and benefits progressively fall towards zero as the Do Something moves closer to the threshold.
- The third phase is when both Do Minimum and Do Something demand are capped as the generalised time are the same. Therefore, there is no difference between the two scenarios and at this point the benefits of implementing the Do Something scenario over the Do Minimum are zero.

Table 10 summarises the PVB for each of the generalised cost cap tests.

Table 10 – Highway PVB for a Generalised Cost-based Demand Cap

Scenario	PVB (£)
No Cap	£6,303
175% GC Cap	£5,931
150% GC Cap	£3,910
125% GC Cap	£1,817
110% GC Cap	£672

Source: Arup

The reduction in scheme benefits on account of the generalised cost cap when compared with the alternatives is explained by the treatment of the demand suppressed by the lack of capacity in the do-minimum scenario. In practice much of this suppressed demand would find alternative times of travel, modes and routes, each choice resulting in a transport user cost. Benefits are reduced because these costs, which count as a benefit of the do-something in the other capping options, are omitted in a much earlier year in the do-minimum than when demand is allowed to grow for each scenario over the same number of years.

The implementation of this type of cap may be difficult as the timing of the cap is sensitive to factors including the existing congestion in the base year and the assumptions of the maximum generalised cost threshold. Changes to these values may impact when in the life of the scheme the demand cap would take place.

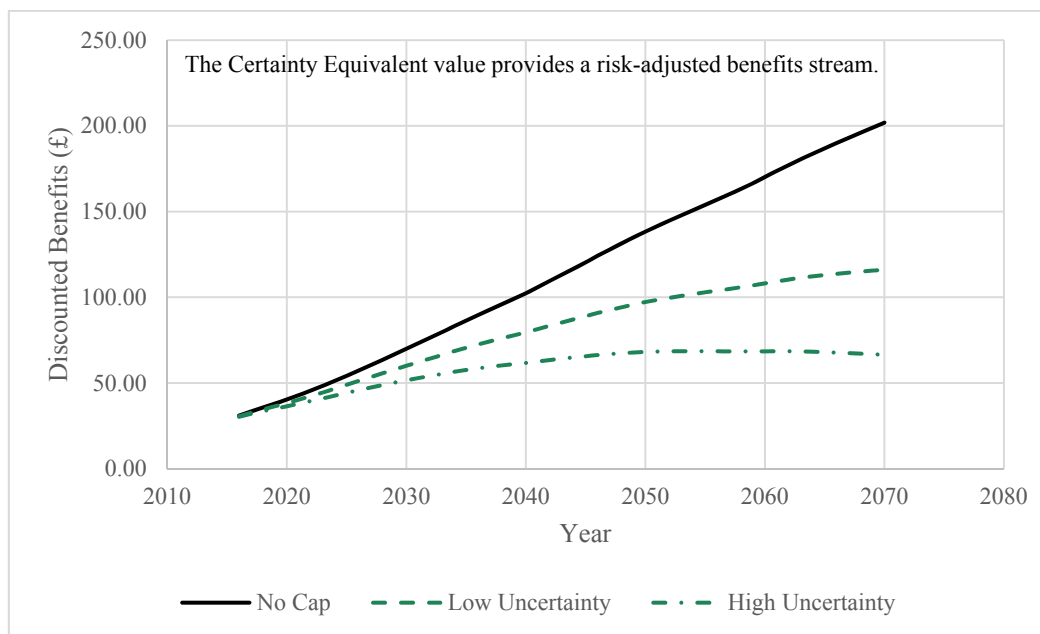
We created a third test to illustrate the change in benefits over time if a certainty equivalent value was applied to the benefits stream. This value is introduced as a potential method to account for issues of uncertainty in demand forecasting (discussed in Section 1.7), including uncertainties in the demand curve, the supply curve, and in future policy. A 1.0% per year reduction was applied to the uncapped benefits for the ‘low uncertainty’ case and a 2.0% per year reduction

was applied to the uncapped benefits for the ‘high uncertainty’ case. For the Low case, the opening year PVB is almost unchanged: the PVB at year 20 is 17% lower and the PVB at the final year element is 42% lower than the uncapped scenario. For the High case, the PVB at year 20 is 32%, and the PVB at the final year is 67% lower than the uncapped scenario.

No attempt has been made here to quantify the appropriate certainty equivalent value, but instead, these tests have been conducted to illustrate the potential effects. In practice, determining the appropriate certainty equivalent value would require an estimation of the potential variance of future demand forecasts and benefits, as well as an assumption of the acceptable amount of risk for variability in long-term forecasts. (A higher certainty equivalent value suggests a more risk-averse approach).

Figure 34 shows the resulting discounted benefits charts. In the figure below, the discounted benefits has a similar shape to the demand overlay graph. In each case, the tests have a lower amount of discounted benefits than the uncapped scenario, but compared to the caps based on years or generalised cost, the discounted benefits continue to increase over the appraisal period. **Table 11** summarises the total present value benefits for each of the tests.

Figure 34 – Highway Discounted Benefits for Certainty Equivalent Test



Source: Arup

Table 11 – Highway PVB for Uncertainty Equivalent Test

Scenario	PVB (£)
No Cap	£6,303
Low Uncertainty	£4,480
High Uncertainty	£3,244

Source: Arup

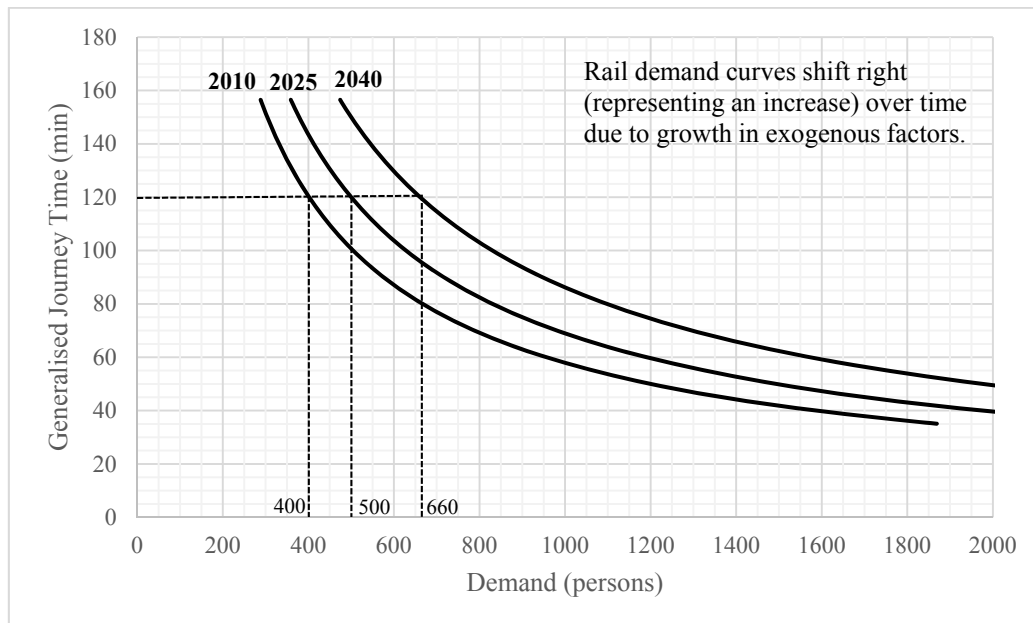
4.1.2 Rail Simulation Model

A separate model was developed for the rail context. It operates in an essentially similar way to the road context, except crowding on trains – instead of congestion – is the limiting factor. This simulation model estimates the demand and supply curves for each year of the appraisal period using guidance within the Passenger Demand Forecasting Handbook (PDFH) v5.1.

In the rail context, the demand curve represents the user response to changes in the cost to use a rail service. It illustrates the peak-hour demand for a train service in a given year over a range of costs. This cost is also known as generalised journey time (GJT) for a specified trip, defined as the sum of in-vehicle journey time, service interval penalty (waiting time), interchange penalty, and a crowding penalty. For this model, we have assumed that demand reacts to changes in GJT with an elasticity consistent for the London and South East area. This means that for higher GJTs, there will be a reduction in demand, while lower GJTs will cause an increase in demand. Note that fares are not a component of GJT, so the equilibrium point on a single demand curve is only affected by changes in travel time and crowding levels.

Based on PDFH, the demand growth over the appraisal period is a function of several exogenous factors including population, GDP, employment, car ownership, fuel costs and fares. As with the road model, WebTAG inputs are used for the change in population, GDP, fuel cost and value of time. As well, national employment forecasts are sourced from the Office of Budget Responsibility (OBR), and national car ownership forecasts are estimated from TEMPRO. For this example, fares are assumed to remain constant (in real terms). Growth in rail demand is illustrated as the demand curve shifting towards the right, which is shown for 2010, 2025, and 2040 in **Figure 35**.

Figure 35 – Example Rail Demand Curves for 2010, 2025 and 2040



Source: Arup

In the above example, if the generalised journey time remained fixed at 120 minutes, then the demand would grow from its base level of 400 persons in 2010 to 500 persons by 2025, and 660 persons by 2040.

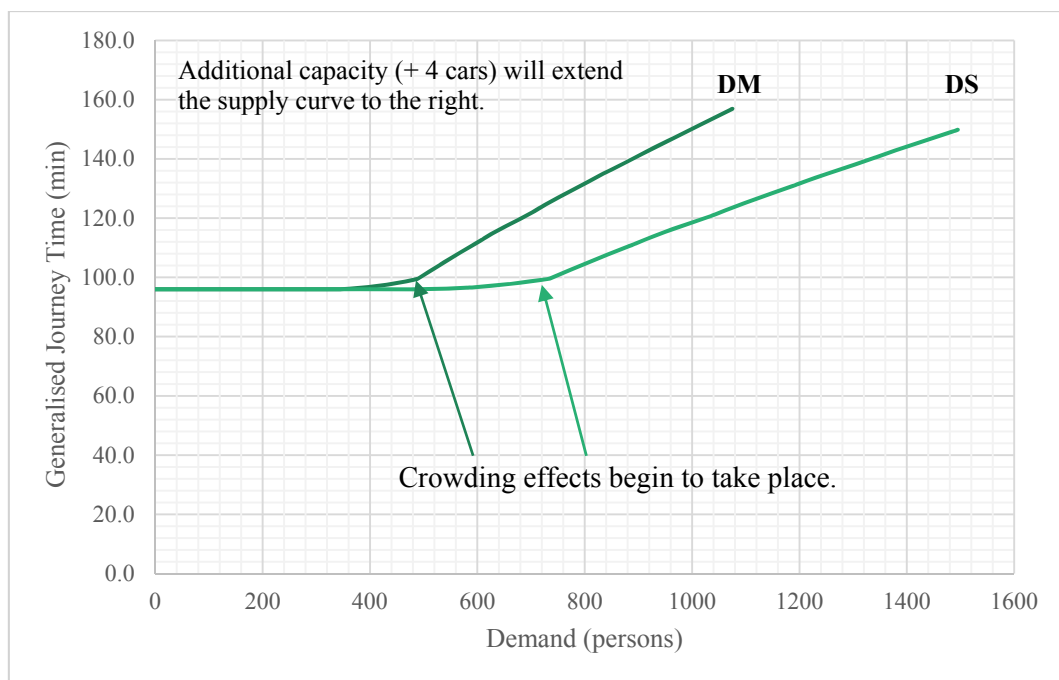
The supply curve in the rail model represents the infrastructure response to demand, and illustrates the GJT to use the rail service over a range of demand, which is the number of passengers on the rail service. Given that all other factors – journey time, service interval penalty, and interchange penalty – are held constant, the only component that affects GJT is the crowding penalty, which increases as the seating capacity and standing capacity of the train becomes occupied with passengers. This model uses both crowding factors and rolling stock capacities from PDFH v5.1.

All components of the supply curve – journey time, frequency and capacity – are within the control of the train operating company (TOC). There are no external factors that can affect the supply curve each year. However, the supply curve may shift due to a proposed service change, such as:

- Improvements to journey time, which will shift the supply curve down;
- Increases to capacity, which will extend the supply curve towards the right, and reduce the slope of the supply curve; or
- An increase in frequency, which combines the effect of the first two improvements.

To illustrate the shift in the supply curve, consider the following simple scenario for a TOC, where the Do Minimum (DM) option is a train service is operated by an 8-car train, while the Do Something (DS) option is operated by a 12-car train. **Figure 36** illustrates the change in the supply curve.

Figure 36 – Rail DM (8-car) and DS (12-car) Supply Curves



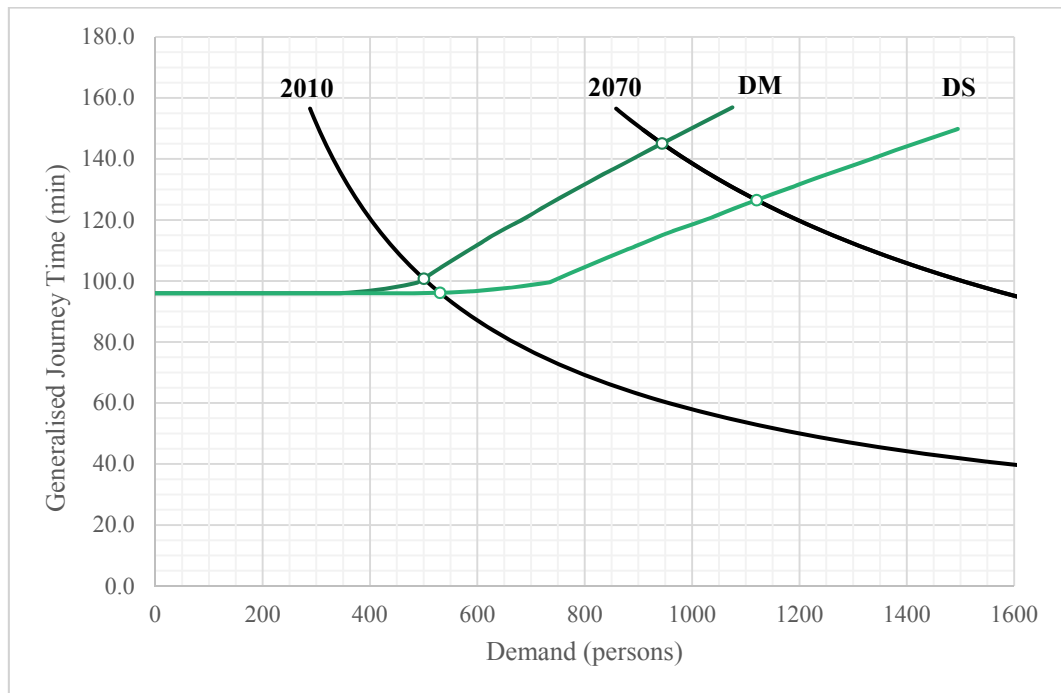
Source: Arup

As expected, the 12-car train will have additional capacity, which causes crowding effects to take place at a higher demand than the Do Minimum option.

Note that this train operator currently operates this as an hourly service to London. The journey time is 60 minutes, and with an hourly service, the service interval penalty is 36 minutes, for a base GJT (without crowding) of 96 minutes. There are no other changes to frequency or journey times.

To calculate the benefits of this scheme, the model then calculates the intersection of the demand and supply curves over the appraisal period. The base year demand in 2010 for the service is 500 persons. For demonstration purposes, the model calculates the demand-supply curve equilibrium and the discounted benefits over a 60 year period (although it is recognised that in practice the appraisal period of a scheme of this scope would be determined by the life of the rolling stock). The supply-demand curves for the Do Minimum and the Do Something case is illustrated in **Figure 37**.

Figure 37 – Change in Rail Demand - Supply Curves and Equilibrium



Source: Arup

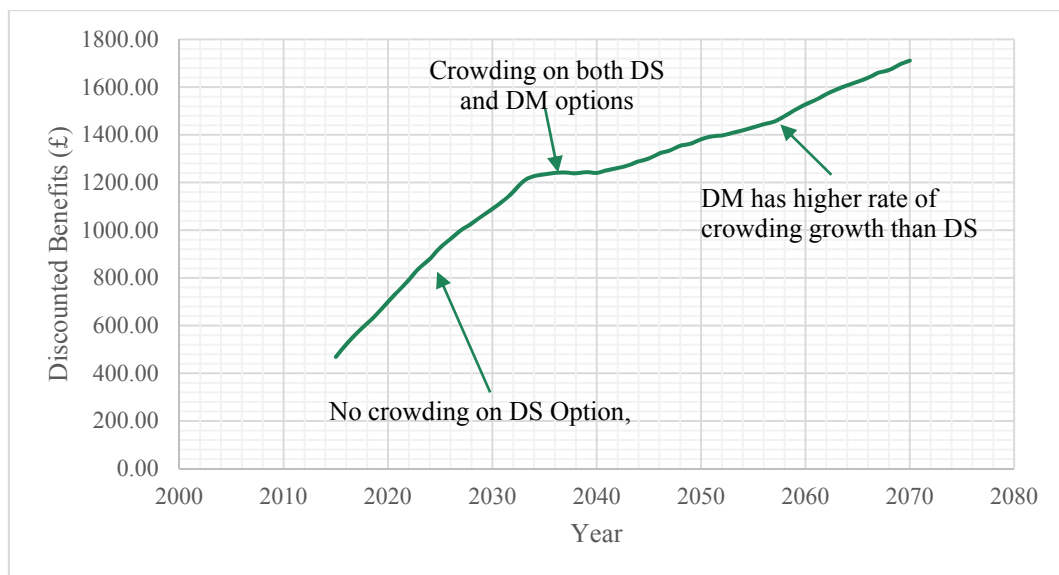
As shown in the above figure, the 2010 Do Minimum equilibrium is at 500 passengers and a GJT of 101 minutes. As the base GJT (without crowding) is 96 min, this means that with the current demand there is some level of crowding. Over the course of a 60 year period, the demand curve will shift right due to changes in exogenous demand factors. The difference in the demand and generalised journey time between the DS and DM options are summarised in **Table 12**.

Table 12 – 2070 DM and DS Equilibrium Demand and Generalised Journey Time

Scenario	Demand (persons)	GJT (min)
Do Minimum (DM)	944	145
Do Something (DS)	1,120	127

Source: Arup

If the TOC continued with its 8-car train, it would carry fewer passengers at a higher GJT. Implementing the proposed scheme would allow the TOC to carry more passengers at a lower GJT. The time savings experienced by each user of the scheme can be converted to a discounted benefit using a value of time and the Rule of Half. The model calculates these benefits for each year of the appraisal period, which is shown in **Figure 38**.

Figure 38 – Rail Discounted Benefits over the Appraisal Period (Uncapped Demand)

Source: Arup

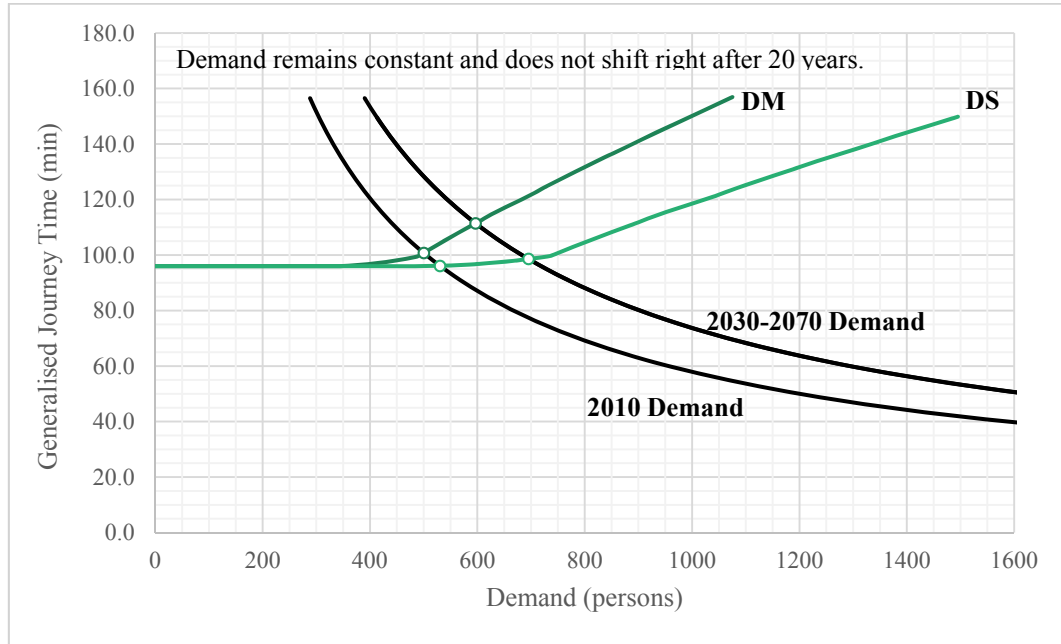
The discounted benefits for this uncapped scenario trend upwards over the course of the appraisal period. The non-linearity of the benefits is due to a couple of reasons. First, the simulation model relies on PDFH factors for crowding, which are presented as discrete factors as opposed to a continuous function. This presents some difficulty for the simulation model to find the exact equilibrium point. While linear interpolation has been undertaken to smooth out the crowding effect, some of the ‘noise’ is still observed in the discounted benefits graph. The second reason is that the points of inflection represent different phases of crowding in the DM and DS options:

- In phase 1: the DM option is experiencing crowding while the DS option remains uncrowded;
- In phase 2: the benefits flattens out in 2032-2040 as the DS option begins to experience crowding; and
- In phase 3: the benefits increase once again in 2040 when the DM option experiences crowding growth at a higher rate than the DS option.

The overall present value benefits [PVB] for this hypothetical scheme is £68,900.

As with the road simulation, the model was used to test the effect of various capping scenarios. A year cap test considered the effect of stopping all demand drivers after 20, 30, 40 and 50 years into the appraisal period. An example demand-supply curve for the 20 year cap is presented in **Figure 39**.

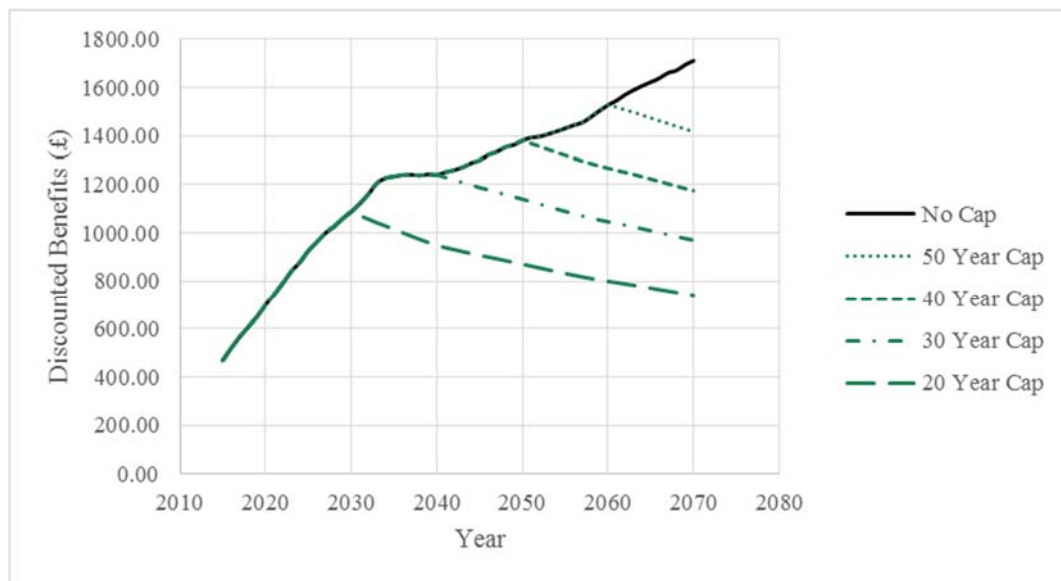
Figure 39 – Rail Demand and Supply Curves for 20-Year Demand Cap



Source: Arup

With the demand-cap based on a year approach, the demand curve is frozen in place after a specified amount of time. The result of this test is in **Figure 40**.

Figure 40 – Rail Discounted Benefits for Years-based Demand Cap



Source: Arup

As in the road model, each scenario of capping follows the ‘No Cap’ scenario until the year the cap is implemented. Once the demand cap is implemented, the discounted benefits decrease. A comparison of the PVB for each cap is provided in **Table 13**.

Note that for rail, revenue is treated as a negative cost, which has the effect of increasing the BCR. In this work, for simplicity, we haven’t modelled the impact of revenue on costs, as the focus of the study is on benefits.

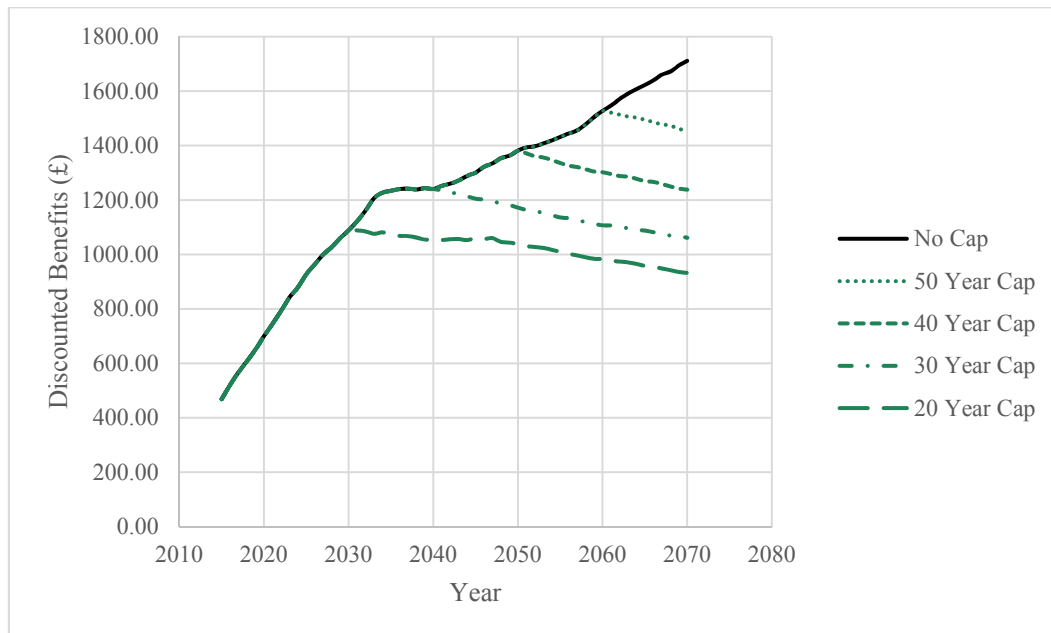
Table 13 – Rail PVB for Demand Cap using Years-based Approach

Scenario	Uncapped Demand PVB (£)	Capped Demand PVB (£)	Total PVB (£)
No Cap	£68,900	--	£68,900
50 Year Cap	£51,060	£16,200	£67,260
40 Year Cap	£36,740	£26,720	£63,460
30 Year Cap	£23,760	£33,990	£57,750
20 Year Cap	£11,750	£36,210	£47,960

Source: Arup

A variant of this test, where all exogenous demand factors except for population (i.e. GDP and employment) are capped, was also simulated in the model. The results of this test, in **Figure 41**, show that more benefits are realised if the population factor continues to influence demand. A comparison of the PVB for each cap is provided in **Table 14**.

Figure 41 – Rail Discounted Benefits for GDP / Employment Years-based Demand Cap



Source: Arup

Table 14 – Rail PVB for GDP / Employment only Years-based Demand Cap

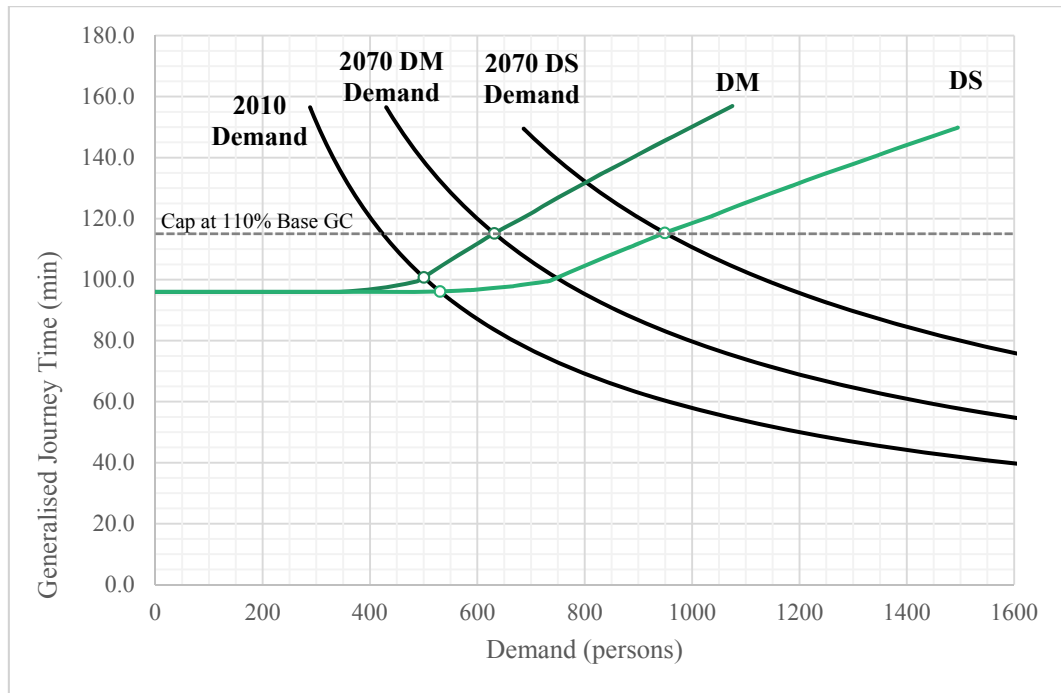
Scenario	Uncapped Demand PVB (£)	Capped Demand PVB (£)	Total PVB (£)
No Cap	£68,900	--	£68,900
50 Year Cap	£51,060	£16,430	£67,490
40 Year Cap	£36,740	£27,400	£64,140
30 Year Cap	£23,760	£35,480	£59,240
20 Year Cap	£11,750	£41,960	£53,710

Source: Arup

In common with the highway schemes, the benefits predicted in the 60th year under these scenarios for capping exceed the benefits in the years immediately after opening. And it follows from this that, if we can be confident about these benefits 60 years hence, the curtailment of the appraisal period to one of 60 years has the effect of omitting a potentially significant source of benefits arising beyond the appraisal period.

The generalised cost cap variant that was tested considered generalised cost (in time units), defined as the sum of generalised journey time, plus the fare (converted to time units using the base year value of time). With the generalised cost cap, the demand would be capped once the generalised cost reaches a certain threshold of the base year generalised cost. In the rail context, under constant fares, journey time and frequencies, this means that the demand would be capped once the crowding exceeded a certain threshold of the base year crowding levels. When both Do Minimum and Do Something demand is capped, the demand levels are at a point where the generalised cost is the same for both scenarios. **Figure 42** illustrates the demand-supply curves for an example this test (110% Base Year GC).

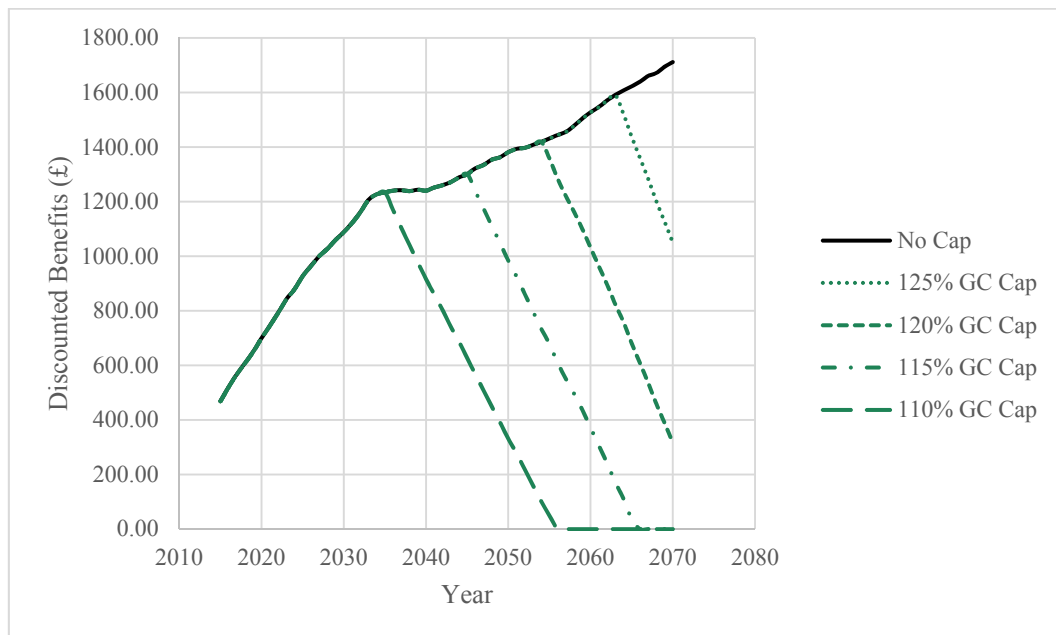
Figure 42 – Change in Rail Demand-Supply Curves with a Generalised Cost Cap



Source: Arup

The Generalised Cost thresholds that were considered in this test include 110%, 115%, 120% and 125% of the base year generalised cost. (Note that these thresholds are illustrative, but in practice will depend on the proportion of the generalised cost that is attributed to fares and crowding). The impact of these thresholds on the discounted benefits is in **Figure 43**.

Figure 43 – Rail Discounted Benefits for Generalised Cost-based Demand Cap



Source: Arup

Similar to the highway model, there are three phases to the rail discounted benefits graph where there is a generalised cost cap. The discounted benefits approach zero when the both DM and DS demand are capped as there are no differences in terms of fares, journey time or frequency, and therefore the generalised cost is based purely on the level of crowding. At the point where both Do Minimum and Do Something generalised costs are equal, the crowding levels are the same in both scenarios.

In common with the roads scenarios, the reduction in benefits when compared with a years based cap is explained by the implicit assumption in the test that passengers who would benefit from the new trains under the DS scheme but are crowded off and unable to travel in the modelling of demand under the DM experience no additional costs. In reality, such passengers would choose a different mode, destination or time of travel, or perhaps continue to travel on trains which exceed the modelled levels of crowding. In each case, they would experience a cost when compared to the DS alternative which would allow them to continue making the trip. A comparison of the PVB for each cap is provided in **Table 15**.

Table 15 – Rail PVB for Generalised Cost-based Demand Cap

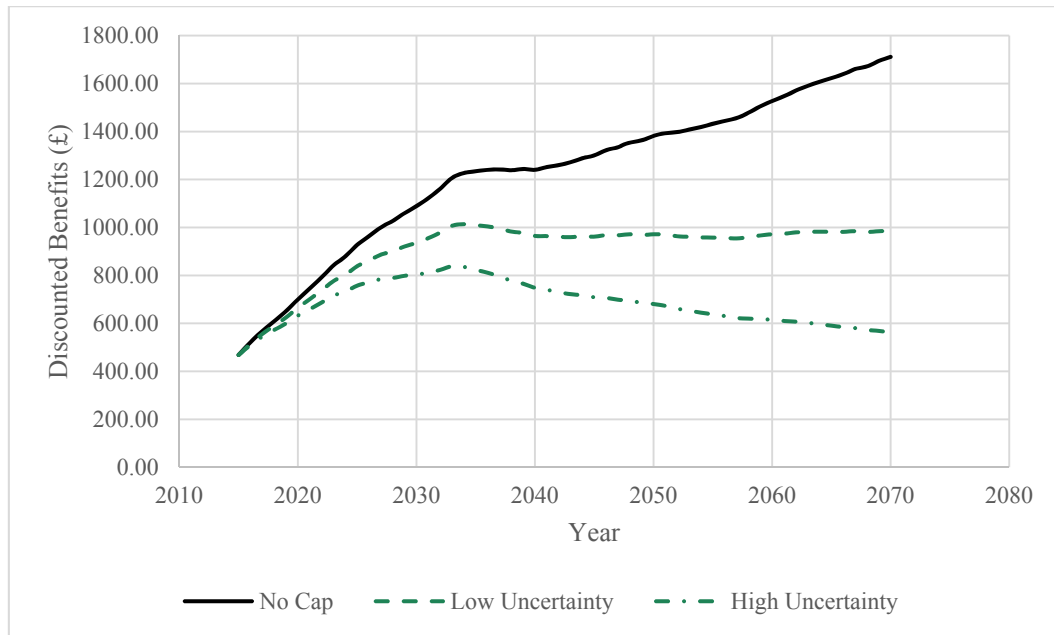
Scenario	PVB (£)
No Cap	£68,900
125% GC Cap	£66,270
120% GC Cap	£57,320
115% GC Cap	£44,240
110% GC Cap	£30,760

Source: Arup

A final test was conducted illustrates the change in benefits a certainty equivalent value was applied to the rail scheme to account for uncertainties in future forecasts. The same two cases (as in the road context) were conducted, where a 1.0% per year reduction was applied to the uncapped benefits for the ‘low uncertainty’ case and a 2.0% per year reduction was applied to the uncapped benefits for the ‘high uncertainty’ case.

The demand overlays were tested in the rail simulation model. The resulting discounted benefits are presented in **Figure 44**. For the ‘Low’ case, the PVB at year 20 is 17% lower, and the PVB at the final year is 42% lower than the uncapped demand case. For the ‘High’ case, the PVB at year 20 is 32% lower, and the PVB at the final year is 67% lower than the uncapped demand case. A comparison of the PVB for this test is presented in **Table 16**.

Figure 44 – Rail Discounted Benefits for Certainty Equivalent Test



Source: Arup

Table 16 – Rail PVB for Uncertainty Equivalent Test

Scenario	PVB (£)
No Cap	£68,900
Low Uncertainty	£50,770
High Uncertainty	£38,220

Source: Arup

4.2 Project Case Study

Appendix B contains a case study of a real project (anonymised) which has had additional strategic model runs far into the future and the benefits and BCRs have been re-calculated. The primary insights here are:

- as expected, the removal of the demand cap resulted in higher BCRs (note in these cases, an uncertainty risk adjustment was not applied); and
- the strategic models, particularly in the roads case, had issues with convergence.

4.3 Conclusions

From the Project Case Study outlined above, as well as practitioner experience, our view is that recommending explicitly modelling demand with a complex transport model over a 60 year period for all projects would not be a sensible proposal. This is due to the instability of transport models far into the future and the time and cost required to develop them which is further discussed in Chapter 5.

The two simulations (road and rail) demonstrate the effect of different scenarios for the demand cap on the change in demand and on the discounted benefits for a simple highway and railway scheme. The tool that was used to conduct these simulations could also be applied to test a wide range of schemes with varying base conditions, rates of growth, demand and supply elasticities, and policy assumptions. But from this work, we conclude that, depending on the circumstances:

- the cap can have a big effect on the present value of the benefits stream relative to the uncapped scenario – see for example the solid line versus the dashed curves in **Figure 40** for rail;
- the formulation of the cap is significant in terms of whether traffic growth due to all sources is capped or whether some sources of growth such as population are uncapped (compare **Figure 40** and **Figure 41**);
- the timing of the cap is significant (compare the alternative dashed curves in **Figure 41**);
- most crucially, whether the cap is conceived of as a ‘year ahead’ concept or as a ‘capacity limit reached’ concept can make a big difference (compare **Figure 40** and **Figure 43**) and we discuss this further below.

The message we take from this exercise is that current guidance has the important virtue of simplicity both in concept and in applicability at scheme level within TUBA. But the particular definition of the cap is essentially arbitrary and the work illustrates that the outcome for BCRs and NPVs is sensitive to definition.

The main line of the guidance focusses on the year ahead form of the cap but the subsidiary line of the guidance asks promoters to have regard to the capacity conditions in the future. We think this is crucial. We stress that the analysis above is illustrative and it is best to regard a minimum level of service or maximum generalised cost threshold as a limiting case. But the more general case where in the latter half of the appraisal period the difference between the generalised cost in DM and DS is significantly eroded by congestion/crowding is, we believe, entirely realistic. Imposing the cap a given number of years ahead and fixing demand at the cap level is probably quite a generous rule where schemes are close to capacity by the cap year.

This raises a number of issues for appraisal some of which go beyond the scope of this report. The first is what is the true economic benefit of increased capacity (i.e. carrying more traffic down a corridor at given generalised cost). The answer is likely to be found in the wider economy impacts such as agglomeration and

labour market effects rather than the direct transport benefits. The second is whether a fixed Reference Case (or Do Minimum) is really a suitable construct for an appraisal period of sixty years. Perhaps the cap is really a construct for saying that the future beyond the modelling period is unknown but policy will be to maintain conditions no worse than in the last year of the model period. The realism of giving effect to that policy assumption can then be debated. The work above suggests that the Department's value for money metrics as they apply to a range of schemes are likely to be quite sensitive to the definition of this construct. The certainty equivalent approach provides an opportunity to do better and we return to this in Chapter 6.

5 Long Term Demand Modelling

This chapter looks at some important considerations for modelling future year demand. In general, there are two methods used to forecast transport demand:

1. strategic models; and
2. elasticity-based models.

This Chapter outlines challenges associated with both of these types of methods in the context of long term demand forecasting.

- **Strategic Models And Cross Sectional Data** – From a theoretical perspective, can we use cross sectional data in strategic models over the long term? There is no conclusive evidence on this.
- **Practical Use of Strategic Models** – This short section describes some of the practical challenges associated with long term forecasting using strategic models.
- **Long Term Use of Elasticities** – This section comments on the use of elasticities over the long term. We found no evidence that elasticities should be used for a particular length of time, however the GDP elasticity has remained consistent over time suggesting its long term use may be valid. (Recent changes in the GDP elasticity in PDFHv5.1 accounts for structural and spatial changes rather than a change in long term GDP elasticity.)

5.1 Strategic Models And Cross Sectional Data

In short, there is no simple and conclusive answer to the question ‘*can we use cross sectional data in strategic models over the long term?*’

In contrast to the forecasting methods for rail (generally elasticity-based models), which derive essentially from time series analysis, multi-modal models, as well as uni-modal models developed for highway appraisal, are typically developed from cross-sectional data. As opposed to longitudinal data, *cross-sectional data* usually refers to data from a particular point in time, and is thus more concerned with the spatial variation in behaviour, as well as that between different segments of the population. An important aspect of these models is to synthesise the base year level of demand for all possible movements – something which is less necessary for rail modelling because of the availability of LENNON.

There thus arises a question as to *whether the sensitivities derived from cross-sectional data are appropriate for forecasting*.

The literature offers limited evidence regarding the stability of cross-sectional relationships over time. The key concept here is that of ‘temporal transferability’, which considers whether ‘*model parameters that best explain travel behaviour at the time at which the estimation data was collected will also explain future travel behaviour*’ (Fox, 2015).

In principle, temporal transferability can be tested by comparing a common model specification estimated on different datasets collected at two or more points in time in the same geographical area. In practice, however, it may not be easy to maintain consistency in the definition and/or recording of ‘common’ data items over time.

An interesting argument advanced in the academic literature (e.g. Ben-Akiva & Atherton, 1977)¹⁵ is that ‘disaggregate’ travel demand models (specified at the level of the individual traveller) may offer greater transferability than more ‘aggregate’ models (specified at the zonal level, for instance).

Given the increasing importance of long-term demand forecasting for policy and planning, Fox’s (2015)¹⁶ thesis represents a timely and authoritative contribution to the literature. From his review of the extant literature, he concludes that: *‘Overall, the direct tests of transferability...are supportive of the hypothesis that mode choice models can be transferred over time, with the majority of studies concluding the models tested were transferable. Furthermore, some of the validation studies demonstrate the models are able to predict the impact on mode share of substantial changes in level-of-service over short periods’*. Fox defines ‘short periods’ as **up to 10 years**.

Informed by his literature review, Fox’s own empirical research then investigates the temporal transferability of mode-destination models over long-term forecasting horizons – which he defines as 20+ years. This investigation is conducted using data from Toronto and Sydney, comparing the transferability of commuter and non-commuter travel separately for a range of metrics (e.g. value of time, cost sensitivity, etc.). Fox finds the **commuter models to be more transferable than the non-commuter models**. Furthermore, the inclusion of socio-economic variables improves model transferability, and travel time and socio-economic parameters are found to be more transferable than cost parameters and alternative-specific constants.

5.2 Practical Use of Strategic Models

Drawing on the collective practitioner experience available to the study team, it is considered not **practical** to recommend running transport models (by which we mean strategic models or multi-modal models) far into the future due to cost to projects, model run times, challenges with defining do-minimum and model convergence. The transport model can become unusable in the context of continuing demand growth. As ever increasing demand is fed on to the networks, the processes in the model which bring demand and supply into equilibrium have an increased likelihood of failing to operate effectively.

For these reasons, we would not recommend requiring practitioners to develop ‘complex models’ beyond 20 years from when their scheme opens. This could be refuted by further research running many models far into the future.

¹⁵ Ben-Akiva, M. & Atherton, T. (1977) Methodology for Short-Range Travel Demand Predictions. Journal of Transport Economics and Policy 11, pp224-261.

¹⁶ Fox, J.F. (2015) Temporal transferability of mode-destination choice models. Unpublished PhD thesis, Institute for Transport Studies, University of Leeds, May 2015.

However, it is reasonable to expect that practitioners can develop forecasts beyond this time by either extrapolating growth from the complex model and/or applying sensible assumptions relating to drivers of long term demand growth. We would recommend, however, that DfT provides guidance on how this should be undertaken.

Note that elasticity-based models are typically implemented in a spreadsheet and can be run easily into the future from a practical point of view. (Note that this point excludes the Planet Framework Model which is primarily a ‘strategic’ model but also utilises elasticities.)

As part of this study a road case study was investigated – which involved long term demand forecasting using strategic models. This is outlined in Appendix B. The case study discusses the limitations with transport modelling far into the future, particularly with convergence in the do-minimum scenario.

5.3 Long Term Use of Elasticities

The key question here is: *‘What is the appropriate time horizon for forecasting with PDFH elasticities?’* First we provide an overview of elasticity based modelling and then we review the timeframe of their use.

5.3.1 Overview

The Passenger Demand Forecasting Handbook (PDFH) details a comprehensive and well-established set of methods, supported by empirical evidence, for forecasting passenger rail demand in Great Britain. The original objective of PDFH was to standardise demand forecasting methods across the post-privatised industry, and the existence of such a resource – at least at this level of detail – makes the PDFH rather unique among national and international railway planning agencies.

In practice, the PDFH framework is typically applied at the scheme level, to forecast the patronage impacts of a given intervention. However, PDFH has also been applied more generally to forecast patronage impacts at the market level (e.g. in the context of DfT’s ongoing Rail Demand Forecasting Estimation (RDFE) study).

The substantive features of the forecasting model are as follows:-

- **Unimodality:** The model is ‘unimodal’, focussing upon patronage change within rail, albeit with some (rather cursory) acknowledgement of the propensity for mode shift.
- **Proportional demand:** In general, the model specifies both dependent (i.e. demand) and independent (i.e. drivers of demand) variables as ratios, where the numerator refers to a ‘do something’ case and the denominator refers to a ‘base’ case.
- **Constant elasticity form:** The model is specified as a log-log function. Given this functional form, the parameters of the model can be readily interpreted as elasticities of demand. Furthermore, the elasticities will be

independent of the magnitude of the independent variables, such that the log-log form is also commonly referred to as the ‘constant elasticity’ form.

- **Standardised explanatory variables:** Methods and evidence are provided to allow the analyst to select – from an exhaustive set of independent variables – those most relevant for forecasting patronage change in the context of a particular scheme. Broadly speaking, these independent variables cover external factors, fare, generalised journey time (GJT), and a raft of quality-related variables.
- **Representation of longitudinal effects *per se*:** The influence of time *per se* on patronage is primarily represented through the ‘external factors’ variable, which includes GDP and population.
- **Representation of quality variables in GJT units:** A key feature of the model is the reliance on generalised journey time (GJT) as the numeraire for quality improvements, where the conversion between GJT and quality is informed by willingness-to-pay (WTP) evidence.

In considering – *at a conceptual level* – the appropriate time horizon for forecasting using the PDFH approach, it would seem sensible to consider any temporal dimensions of the substantive features above.

- **Unimodality:** Whilst the model takes account of the cost, quality and availability of other modes at a fairly coarse level, it is questionable whether this representation properly accounts for any substantive change in the attractiveness of these modes over time. There is also the particular challenge of representing situations where new modes emerge, which is obviously more of a possibility the longer the forecasting profile.
- **Proportional demand:** An issue here is in defining the ‘base’ case (i.e. the denominator of the ratios, for both dependent and independent variables). In the context of PDFH, the base case is generally interpreted as a ‘*known level of base demand between stations, on a route or at some more aggregate regional, train operating company or national level*’ (PDFH v5.1). However, if demand forecasts are to be applied in appraisal, then the focus of interest will be upon patronage change relative to a do minimum (DM), and this DM may be difficult to conceptualise over the long term.
- **Constant elasticity form:** Whilst the CE form brings a number of practical attractions, a key issue in the context of long-term forecasting is the constraint that the elasticity will remain constant irrespective of the duration of the forecasting profile. In this regard, the logic of reassurance offered by PDFH would seem dubious: ‘*Generally the position taken in the PDFH is that the same elasticity (or strictly the same functional form) applies to small or large changes in variables. This must logically be true in the long term; as the identity of the people making the trip changes, the history of how variables reached their current levels becomes irrelevant*’ (PDFH, v5.1). In contrast to the CE form, more flexible dynamic econometric forms used widely in the econometrics literature allow the estimation of both short run and long run elasticities in combination. In fairness, PDFH has actively engaged with some of the issues in this area; for example, PDFH cites evidence from dynamic econometric rail demand studies, and research has been commissioned on the

specific issue of ‘large’ changes in GJT, although this found ‘*no support for size or sign effects on the GJT elasticity*’ (PDFH, v5.1). Despite some ambiguities in the underlying evidence base, a recent development in PDFH (in moving from v5 to v5.1) has been to distinguish between ‘short’ and ‘long’ term GDP elasticities for non-London flows. However, PDFH remains less than definitive on its distinction between ‘short’ and ‘long’: ‘*There is no objective evidence as to when the longer term elasticities should be used (or indeed that the elasticities should reduce), but we would recommend that the short term would be for the next five years, long term after ten years, with a gradual transition between them*’ (PDFH, v5.1).

- **Standardised explanatory variables:** An issue here is whether the ‘standard’ set of variables agreed at a given point in time will remain representative of the key drivers of demand many years hence. This would seem less problematic for external factors, fares and GJT, but more problematic for some of the quality variables, especially where there is a technological dimension (e.g. availability of wi-fi) which has the potential to stimulate non-trivial changes in travel behaviour.
- **Representation of longitudinal effects *per se*:** In the mid and late 1990s, the PDFH forecasting model included an explicit (negative) time trend alongside the (positive) GDP elasticity, which was designed to reflect observed changes in car ownership, road building and new car journey opportunities, lower fuel and operating costs and increased coach competition. More recently, the time trend has been removed, such that GDP – which, in the absence of recession, would be expected to increase over time – now represents the primary indicator of temporal growth in patronage *per se*. With regards to the level of spatial aggregation of GDP data, PDFH v5.1 recommends that: ‘*The most suitable income measure is that at the most local level. Failing that, forecasts of regional GDP or GVA per capita can be used. The default is to use national growth estimates*’.
- **Representation of quality variables in GJT units:** A fundamental feature of the PDFH forecasting framework is the proposition that, informed by WTP evidence, changes in quality can be readily converted into GJT units, such that the influence of these quality changes on patronage can be analysed through the numeraire of GJT. For any given quality change, the forecasted patronage response will, all else equal, be driven by a combination of the GJT elasticity and the multiplier that converts WTP for the relevant quality attribute into GJT. In this context, the accuracy of long-term forecasts will be dependent not only upon the accuracy of the GJT elasticity and the relevant multipliers, but also upon the aforementioned proposition that changes in quality necessarily flow through into patronage change (given that the GJT elasticity is non-zero). An alternative approach – which could yield quite different forecasts – would be to estimate dedicated elasticities for each of the quality variables. However, even if one subscribes to the use of a numeraire for quality attributes, there is the question of whether time is more appropriate than money for this purposes. From a policy perspective, it would be easier to implement demand management through fares rather than through GJT. A further point to mention is that, whereas the value of money-based variables (e.g. fare) will naturally increase over time with price and income growth, the

value of time-based variables will not, and this could induce bias in their relative values (with implications for demand forecasts) over the long-term.

5.3.2 Appropriate Time Horizon for PDFH Elasticities

As noted in the Phase 1 report of the ongoing DfT ‘Rail Demand Forecasting and Estimation’ (RDFE) study, PDFH substantially underestimated rail demand growth through the mid- and late- 1990s. At that time, PDFH (v3) employed a (positive) GDP elasticity in combination with a (negative) time trend, where the latter was rationalised in terms of the following developments in travel demand prior to the 1990s:

- strong growth in car ownership;
- road building and the advent of new car journey opportunities;
- lower fuel and operating costs; and
- increased competition from coach on inter-city routes.

However, these developments did not in practice continue through the 1990s, and instead travel demand was subject to:

- slowing of car ownership growth;
- congestion increasing and more widespread;
- fuel price increases (especially through the fuel duty escalator); and
- other positive trends for rail (e.g. road unreliability, increased environmental concerns etc.).

Since around 2005 onwards, PDFH has experienced a further period of substantially underestimating rail demand growth (although the level of underestimation has shown variability by market segment). However, this has been against a backdrop of material changes in PDFH method (v4 onwards), notably:

- removal of the time trend; and
- introduction of within-rail cross-elasticities.¹⁷

As part of Phase 1 of the RDFE study, a PDFH ‘backcasting’ exercise was conducted in order to re-examine the relationship between rail demand growth and exogenous drivers, given the elasticities prescribed in PDFH/WebTAG. This exercise was based upon:

- annual data for the period 1998-2014; and
- a simplified version of current WebTAG/PDFH recommendations (combination of directions, no relative population, no ticket type cross-elasticities).

¹⁷ Versions 1 and 2 of PDFH featured both journey time and fare elasticities, whilst version 3 saw the introduction of quality dimensions of the journey, e.g. interchange and crowding. Version 4 further introduced cross-elasticities with other modes, and additional quality dimensions.

Visual inspection of backcasts from this exercise would seem to suggest that:

- PDFH accurately forecasts rail demand growth in 2007 for all market segments;
- PDFH underestimates demand growth from 2007 onwards, again for all segments; and
- PDFH tends to overestimate demand growth prior to 2007, but some segments (e.g. seasons) show exception to this rule.

The latter finding would seem potentially significant, given that contemporaneous forecasts (using a PDFH framework including a time trend) *underestimated* demand growth through the early 2000s.

If the PDFH shows ‘patchiness’ in its ability to accurately estimate demand growth when using fully accurate input data (i.e. as in the case of the backcasting exercise) then its ability to accurately forecast into the future using input data subject to uncertainty is challenging. Against this background, it is worth noting that the ***PDFH-recommended GDP elasticity has remained broadly unchanged through this period***, possibly indicating that forecasting error may be due to quality variables and/or competition with other modes rather than to external factors.

However, acknowledging that PDFH forecasts arise from a combination of data inputs and recommended parameters, it would seem possible in principle to make progress towards understanding the uncertainty inherent within these forecasts by employing an approach similar to that used by DfT to understand uncertainty in their 2012 VoT forecasts (Wheat, Wardman & Bates, 2012)¹⁸. The approach would involve three basic steps, outlined below.

1. Analysing the distributional characteristics of ‘inputs’ (e.g. in the case of PDFH, forecasts of GJT, fares, external factors, quality improvements, etc.) and ‘parameters’ (e.g. in the case of PDFH, elasticities and multipliers that convert quality changes into GJT units), where the former reflects inherent uncertainty in model inputs, whilst the latter reflects uncertainty in parameters which could potentially be reduced by collecting more and/or better quality data.
2. Analysing the manner in which model input/parameter uncertainty is propagated by the model into model ‘output’ uncertainty.
3. Analysing the distributional characteristics of the output uncertainty, for given inputs and parameters.

One practical problem is that PDFH’s parameters arise from several different empirical studies (contrasting with DfT’s model for forecasting 2012 VoT, which derived from only two empirical studies), and this would add complication to the above approach.

¹⁸ Wheat, P., Wardman, M. and Bates, J., Institute for Transport Studies, University of Leeds (2012) Advice on Statistical Confidence of Appraisal Values of Non-Work Time.

6 Discussion

Chapter 1, outlined background to the demand cap and the long term quantification of economic benefits. It specified a number of questions that are key to considering alternative methods for quantifying long term benefits in appraisal including:

1. the ‘**true**’ profile of demand beyond 20 years;
2. the nature of **uncertainty**; and
3. **capacity**, congestion and crowding.

Chapters 2 to 5 present work undertaken to help understand these questions. They summarise research into historic trends, individual scheme profiles, supply-demand equilibrium simulation modelling and issues relating to long term transport demand forecasting.

Bearing in mind the work undertaken so far, this Chapter attempts to discuss the basis of the demand cap, and summarises our proposal for DfT to consider.

6.1 Is the demand cap a valid concept in a literal sense?

There is little or no evidence to support the existence of a sharp cut off or saturation point twenty years into the future. The literal interpretation of a cap as representing some mechanism whereby realised demand growth suddenly plateaus at some fixed future date does not seem to be based on theory or evidence.

From first principles, we would expect realised demand growth to be driven by exogenous factors such as population growth and real income per capita growth, by changes in economic, social and spatial structure, by changes in tastes, preferences, technology, and by changes in the generalised cost of travel. At the scheme level, changes in generalised cost may be partly exogenous (e.g. changes in fuel prices, regulated rail fares) but are partly endogenously driven by the interaction between latent demand growth and increased congestion/crowding which drives up generalised cost and acts as a brake on realised demand growth. These mechanisms are illustrated in the simulation exercise in Chapter 4.

It follows from this that the future realised growth profile question is usefully decomposed into a series of sub-questions.

- What do we expect the trajectory of GDP/capita and population to be?
- How do we expect the GDP elasticity of demand for travel to change as GDP/cap increases?
- How do we expect structural and behavioural trends to impact on the travel market?
- How do we expect generalised costs of travel to develop over time?

- What other supply-side measures, such as capacity constraints, do we expect to carry out in the second half of the appraisal period?
- How do we expect transport policy to respond to the trajectory of supply and demand over time and with what consequences for transport users?
- How do we answer these questions at the correct spatial, economic and social level of granularity required for the appraisal of schemes in particular locations or modes?

These questions go far beyond the scope of this project and in some cases are the focus of other research realised. But we want to note that a demand cap twenty years ahead is not a correct literal interpretation of the way these forces are likely to play out. However, despite not being a correct ‘literal interpretation’ of these forces, the demand cap in practice was a reasonable, simple method of accounting for uncertainty.

6.2 Is a demand cap a fair simple proxy for a complex reality?

Our interpretation of the DfT’s rationale for creating the demand cap in revised guidance in 2004 is as follows. The revised Green Book had changed the appraisal period to sixty years and the discount rate to 3.5% for the next thirty years and 3% beyond that. In typical appraisals, demand would be growing at about the rate of GDP and key values such as the value of time would also be growing directly proportionately to GDP. The modelling guidance suggested modelling the opening year and a design year such as twenty years away from today, with interpolation of benefits between the two dates and extrapolation beyond.

An implication of the Green Book guidance was that in the absence of intervention, a high proportion of the benefits would be accruing far into the future and on the basis of extrapolation of the benefits in the modelled period. This did not accord with DfT’s normal policy and robustness in appraisal. But the solution of extending the modelling period further into the future was also unappealing and burdensome. Therefore DfT chose a relatively simple solution of freezing demand at the twenty year point but allowing unit benefits such as the value of time to continue to grow throughout the appraisal period. This solution was manageable and was argued to be equitable among schemes, scaling down the BCRs in a similar way across schemes relative to the uncapped scenario. It was also defensible at public inquiry; at least we are unaware of the cap coming under attack from scheme objectors. As a device to deal with a problem, the demand cap has certain pragmatic virtues.

Unfortunately, though, the demand cap does not capture the complex reality of demand and uncertainty for the following reasons.

First, we do not think the evidence (in Chapter 3) supports a sudden cut-off of a saturation kind at any particular future date. Secondly, the simulation exercises (in Chapter 4) show that the cap does not treat all schemes equally. For example, schemes for which there is adequate capacity to cope with demand far into the

future have their benefit growth attenuated, while schemes which are close to capacity by year twenty may be quite generously treated by capping demand at year twenty levels. **Figure 43** illustrates the benefits for cases where the demand reaches a capacity limit. Under the scenario where the capacity limit is assumed to be either 110% or 115% of the base year generalised cost, the demand reaches this limit prior to the end of the appraisal period. Therefore, the rule ‘cap after twenty years’ produces higher benefits than if demand was limited by capacity. A comparison of the PVB in **Table 14** and **Table 15** shows that the ‘cap after twenty years’ would be higher under this example. This is a specific example of a generic appraisal question – how best to represent reality in conditions which are far removed from observed conditions.

6.3 How should we deal with uncertainty?

The reality is that there are manifold sources of uncertainty in transport forecasting and appraisal. We do not know how GDP per capita and population will grow, nor do we know how the elasticities to travel demand will evolve, nor the technology and behavioural futures, nor the way in which supply and demand will interact at generalised costs which are very different from current. Nor can we be sure that the values of travel time and other benefits will turn out to have the assumed relationship with income.

Also, transport appraisal takes place under a construct – the reference case is some version of a Do-Minimum scenario including known commitments and the Do – Something is the same scenario with the scheme added to the network. This may be viewed as legitimate for the first third or half of the appraisal period but whether a fixed Reference Case for the whole of a sixty year period is realistic is a serious question. We note that several of the schemes for which evidence is reported in Chapter 3 have experienced further increases in capacity part way through a sixty year appraisal period.

The implication of this is that the central case scenario which is used to generate the scheme benefits as the difference between ‘Do-Something’ and ‘Do-Minimum’ is a prediction within a range which inevitably widens the further out from the present we go. Whether the central case indeed represents the expected (mean) value of the future benefits is one issue. But the more important one is that the appraisal numbers are a single valued estimate of a distribution of outcomes.

Is society indifferent between one option yielding a certain £100 million and another yielding a 50% chance of £50 million and a 50% chance of £150 million? Our view is that it should not be: more to the point the Green Book (Annex 4 para 35) states:

‘A decision maker who is risk averse cares about this potential variability in outcomes and is willing to pay a sum in exchange for certainty (or is willing to put up with variability on receipt of compensation). This compensation is the cost of variability and should be included in appraisal when it is considered appropriate.’

Our view is that there is a strong case for considering a Certainty Equivalence approach to be appropriate when dealing with long-dated infrastructure projects. Such an approach would in our view be a conceptually better way of dealing with the uncertainty phenomena.

6.4 Our Proposal

In this section we give a brief overview of the components of a strategy for handling the benefit estimation for transport sector projects over a sixty year appraisal period. We believe it's important to avoid unnecessarily increasing the modelling and appraisal burden at the individual scheme level. We think that the issues of ensuring that a transport model reaches a stable equilibrium with continuing growth in demand and a constrained network are too great to justify extending the modelled operating period beyond twenty years. However the corollary of this is that work would be required by DfT to propose an approach that would be suitable for general usage.

The approach we propose is outlined below. It is intended as an integrated package rather than a menu to choose from.

- a) A distinction should be drawn between the model period and the extrapolation period. ***We propose that the model period should be the planning and construction period plus a twenty year operating period.*** Typically demand and benefit modelling and appraisal would be undertaken for the opening full year and for the last year of the model period. We would encourage the addition of one or more intermediate years so as to give a better handle on the trajectory of benefits during the operating period and help to inform the diagnostics for extrapolation. Special consideration would need to be given to staged construction schemes. Interpolation rules would be used as now to obtain the complete profile within the model period.
- b) From its existing work on Monte Carlo simulation and fan diagrams, DfT would undertake work on estimating the variance around the expected value of benefits.
- c) In discussion with Treasury and other parties DfT would establish an assumed cost of variability or certainty equivalence using a formula such as that in para 37 of annex 4 of the Green Book (or the version outlined below and described in section 1.7). We suggest this might be implemented as a compound factor over time, making the simple assumption that variance increases proportionately with time.

$$U(X_c) = E[U(X|\sigma^2)]$$

- d) This adjustment factor would be applied over *the entire appraisal period*, from year 1 to year 60. Within the model period, specific guidance exists on relevant growth assumptions and these are built in to relevant worksheets. The appraisal worksheets would require amendment to build in the certainty adjustment factor in the final benefit calculations.

- e) For the extrapolation period, it is important at this stage not to pretend that we know more than we do about conditions in the second half of the appraisal period. So we prefer to see a relatively simple approach, strongly controlled by DfT through its guidance. Essentially DfT will need to consider how the growth fan should be assumed to behave beyond the model period; whether and how capacity considerations and demand/supply interaction effects will apply; other policy considerations such as what future schemes or policy interventions might be relevant. All of these considerations would inform the way in which the trajectory of demand and benefits would be extrapolated from the modelled period to the full appraisal period. One possible way of delivering this *certainty equivalent risk-adjustment* would be to produce a look up table which practitioners would use to move from the model period to the full appraisal period.
- f) We have considered two options extrapolating demand beyond 20 years which we call Option A and Option B.
- g) **Option A** – extrapolate demand from the 20-year complex modelling period based on the model trend but to a capacity cap that may be related to a level of service. We think that further guidance should be developed for practitioners that defines appropriate measures of reaching capacity.
- h) **Option B** – a set of multipliers contained within a workbook in WebTAG. It is premature to design the workbook, but in outline, its function would be to adjust the discounted benefit over the twenty year modelled period to reflect the benefits over the full sixty years. The details of doing this would need to be discussed inside DfT and with interested parties but the design is likely to need to take account of factors such as:
- the trajectory of benefits over the modelled period;
 - the relevant income and population elasticities which might be expected to vary at least with mode;
 - the ratio of demand to capacity in year 20 because this affects the outcome growth in demand over the total appraisal period;
 - the proportionate increase in capacity afforded by the scheme because this, together with the capacity utilisation rate at year 20 affects how long the benefits may hold out; and
 - agreed workable definitions of capacity relevant to urban/inter-urban schemes and different modes.
- i) On this basis, in our proposal, the demand cap in the sense of a limit reached at some particular future date across all schemes regardless of circumstance would then be removed.

6.5 Peer Review

By John Preston, University of Southampton

This report outlines clearly the practical rationale, such as it is, for implementing a demand cap. It also illustrates clearly the consequences of such a demand cap for situations where demand and supply are changing over time. It has a particular strength in putting together data on the long run performance of built schemes which demonstrate that demand continues to grow after 20 years (although in this case this is 20 years after opening, whereas for WebTAG it is 20 years after appraisal and more like 15 years after opening). However, there is some indication of a slow-down in growth after 20 years compared to before, at least for road schemes (**Table 4**) – possibly indicating some tendency towards capacities being reached. Having said that, it is clear that the empirical evidence does not support a sudden cut-off at a particular future date, whilst a 20 years cap cannot be justified theoretically.

For example, the simulation exercises (Chapter 4) illustrate that the cap does not treat all schemes equally. Schemes for which there is adequate capacity to cope with demand far into the future will have their benefits much reduced, whilst schemes that are close to their capacity limit at 20 years may be more generously treated by demand capping. A 20 years demand cap compared to no cap reduces the PVB of the rail scheme by 30% (**Table 13**) but reduces the PVB of the road scheme by 52% (**Table 8**). Therefore, there could be a danger of long lived assets being designed to reach capacity at around the 20 years mark – which is unlikely to be cost effective in practice if additional capacity subsequently has to be provided. The simulations also illustrate the impacts of capping growth related to rising incomes whilst still permitting growth related to rising population levels. They also illustrate the impact of a Generalised Cost based cap and the application of a Certainty Equivalent test, although in both cases there are challenges in making these approaches operational in terms of determining the maximum permitted increase in Generalised Costs or the per annum reduction in benefits.

The proposed distinction between a (complex) model period and an extrapolation (or more strategic model) period is appropriate, as is the recommendation that the model period should be the planning and construction period plus a twenty year operating period, although provisions would need to be made for cases where opening is delayed. Adopting a certainty equivalent risk adjustment has the advantage of being consistent with the Green Book but the key assumption that variance increases proportionately with time would need to be tested, although prima facie this assumption does not seem unreasonable. The options for modelling demand beyond year 20 would need further development in terms of the two specified options (and other options – including a hybrid of the two options proposed). A capacity cap that is related to level of service has some appeal but there is a challenge in determining this level for different schemes.

7 Recommendations

This final chapter summarises three recommendations **for consideration and further research** by DfT drawing on the findings of this study. We see these three recommendations as being considered together rather than as options that can be selected individually.

It's also important to note that **we are not advocating that these recommendations be adopted right now**. More work is required by DfT to assess the impact of these recommendations – particularly to assess whether it's realistic to be able to quantify and then implement the certainty equivalent. We have put forward a proposal of what an alternative to the demand cap could look like.

7.1 20 Years Modelling from Opening

The current approach to demand capping for road, rail and HS2 was outlined in section 1.3 Demand Cap.

For our first recommendation, we recommend that DfT consider specifying that the 20 year complex modelling period be run *from the predicted year of scheme opening* rather than 20 years from the date of the appraisal. This is a matter of judgment based on the consideration of three key issues, which are summarised below.

In summary, for the 20-year complex modelling recommendation, there is a trade-off between 'fairness between schemes' and 'greater uncertainty'. Our view is that for schemes that take longer to appraise and design, it is 'fairer' to explicitly model demand 20 years from opening, and that this outweighs issues of greater uncertainty and modelling practicalities. While there may be greater uncertainty in model parameters and forecasts further into the future, this can be dealt with through the use of certainty equivalence. Also, a 20 year modelling period from opening is deliverable within modelling practicalities, and serves other purposes such as the concept of a "design year".

1. **Fairness.** The concept of fairness between schemes and across modes is important and we believe that schemes that add large capacity are disadvantaged if not explicitly modelled well into future. It is fairer for complex modelling to be undertaken on a *consistent basis for rail and road*, and for big and small schemes.
2. **Uncertainty.** It is true that uncertainty is greater the further you forecast into the future, however this can be accounted for with a certainty equivalence and through other DfT approaches to uncertainty such as the scenario analysis used in the HS2 business case.
3. **Other practicalities.** There are strong practical reasons for this recommendation. In general, when designing the specification of a scheme, we need to determine the size of stations, capacity of the fleet or characteristics of pavement, and for many schemes it has been customary to undertake detailed demand modelling for at least 20 years after scheme

opening. If including the costs of fleet size and stations in the appraisal, we should also be including the benefits that will be obtained from this and explicitly capturing this. If a scheme promoter is going to the effort of designing and building something that is a significant cost, explicitly modelling demand for the first 20 years of its life is entirely reasonable.

Consistency

We believe we're recommending *more consistency* in appraisal rather than less consistency, in particular in recommending a common approach for road and rail. We recognise that this might require more explicit treatment of uncertainty than the present demand cap approach. However, we think it's important that schemes that add more to capacity than the average schemes, despite having a more distant opening date, have this capacity is recognised in the appraisal.

Limitations

We acknowledge that a limitation of this approach is that projected scheme opening dates change, which can change the BCR but the decision should be made based on best estimate of the opening scheme date.

Summary

The 20 year period is a result of a pragmatic judgement based on practical considerations. Undertaking complex modelling 20 years from opening provides a good compromise between the ultimate aim of modelling as far into the future as practical, and the limitations of most transport models. It also reflects a traditional approach to a "design year" which can be used to determine service levels, purchase rolling stock and design other operations.

7.2 Adopt a 'Certainty Equivalence'

Adopting a certainty equivalence involves attaching a higher weighting to a downside outcome in appraisal than to a central outcome. It also has the effect of reducing the weight put on the expected long term outcome and so reduces the likelihood of benefits increasing every successive year until the end of the appraisal period. It effectively increases the discount rate.

1. The Treasury Green Book says you may wish to consider Certainty Equivalence if the uncertainty affects all schemes.
2. To some extent, the certainty equivalence, addresses the issues about what you assume for benefits between years 20 and 60.
3. It has the potential to stop benefits growing into infinity, because then every scheme would have an infinite benefit. The certainty equivalence could therefore address the problem we face under the current approach to capping and under some of the options we have considered of the annual discounted benefits being significant in the last year of the appraisal period and then dropping to zero in the following year.

Attaching more weight to a downside outcome than to a central value has the effect of reducing the weight put on the expected long term outcome and so

reduces the likelihood of benefits increasing every successive year until the end of the appraisal period at year 60. This would replace the previous role that the demand cap had in accounting for the increased uncertainty of long term benefits.

7.3 Remove the Demand Cap

The demand cap is a simple way of taking account of a set of complex phenomena – particularly uncertainty in the longer term – however the rationale was also partly pragmatic. We propose that DfT consider the removal of the demand cap and undertake research on how an uncertainty equivalent could capture uncertainty and be implemented in appraisal.

Part of this research would also involve the development of **guidance** for practitioners around the method for quantifying long term benefits beyond the initial 20-year period as a result of the removal of the demand cap. We propose two options.

- **Option A** – Extrapolate demand from the 20-year complex modelling period based on the model trend but to a capacity cap that may be related to a level of service. We suggest that further guidance should be developed for practitioners that defines appropriate measures of reaching capacity. If this approach were to be adopted, we think that further guidance should be developed for practitioners that defines appropriate measures of reaching capacity.
- **Option B** – Guidance could be developed that involved a set of factors or multipliers which would be used to estimate benefits in the 20-60 year period based on the benefits in the 0-20 year period. Further work would be needed to determine the dimensions of these multipliers but we would expect the multipliers to account for the following:
 - the trajectory of benefits over the 20 year modelled appraisal period;
 - evidence of the saturation of demand up to and beyond year 20, with reference to the National Road Traffic Forecasts (NRTF), population projections etc.;
 - capacity constraints; and
 - the uncertainty about the long term future and the certainty equivalent value of the benefits.

We don't have a firm conclusion on how to estimate or implement the factors (or extrapolation), because we did not have access to a wide range of projects and models but this is something that DfT would need to investigate in order to develop these factors.

Appendix A

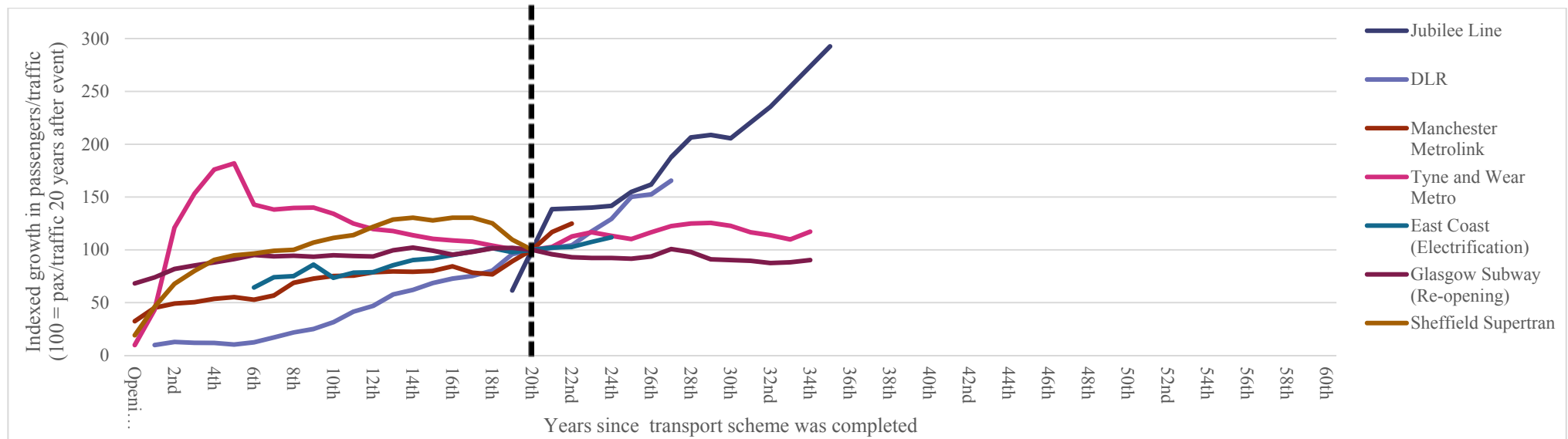
Demand Case Studies

Additional Summary Charts

Figure 45, Figure 46 and Figure 47 below show the actual growth profile of individual transport schemes over time. The figures are indexed to the 20th year after opening in order to demonstrate the difference in growth trajectory between early and later years.

Figure 45, shows the demand profile for sampled rail schemes specifically. Six out of the seven rail schemes sampled grew after the 20th year, although the rate of growth varies substantially. Some variability is probably due to further investment in schemes post-original implementation.

Figure 45 – Growth in passengers for individual rail transport schemes indexed to the 20th year after opening¹⁹

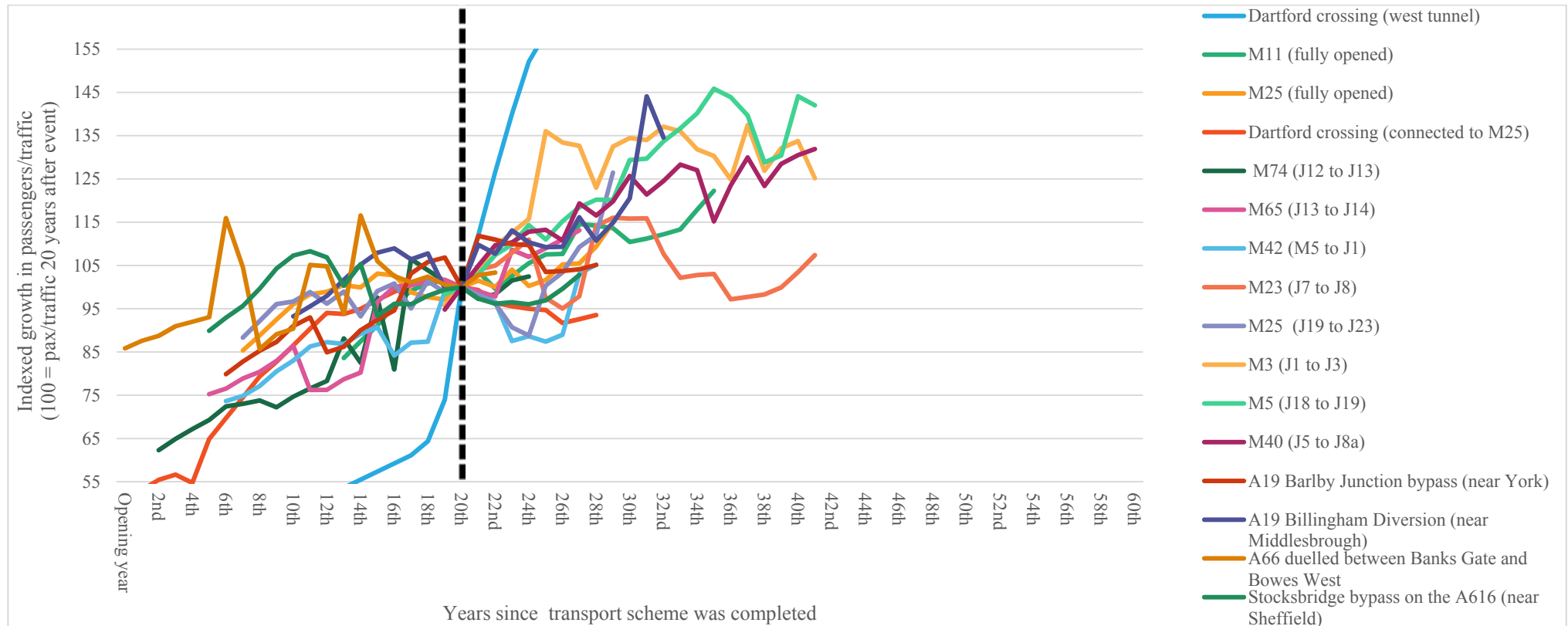


Source: DfT & Rail Industry Monitor

¹⁹ 20th year data was not available from Eurostar, London Underground Victoria Line and Blackpool tram.

Figure 46, shows the demand profile for sampled road schemes specifically. The majority of road schemes were still growing 20 years after opening, although the rate of growth varies substantially. Road schemes demonstrate less of a ramp up in initial years of opening when compared to rail projects, this is probably due to rail schemes tending to deliver larger capacity increases than road projects, which tend to be more incremental capacity increases.

Figure 46 – Growth in passengers for individual road transport schemes indexed to the 20th year after opening

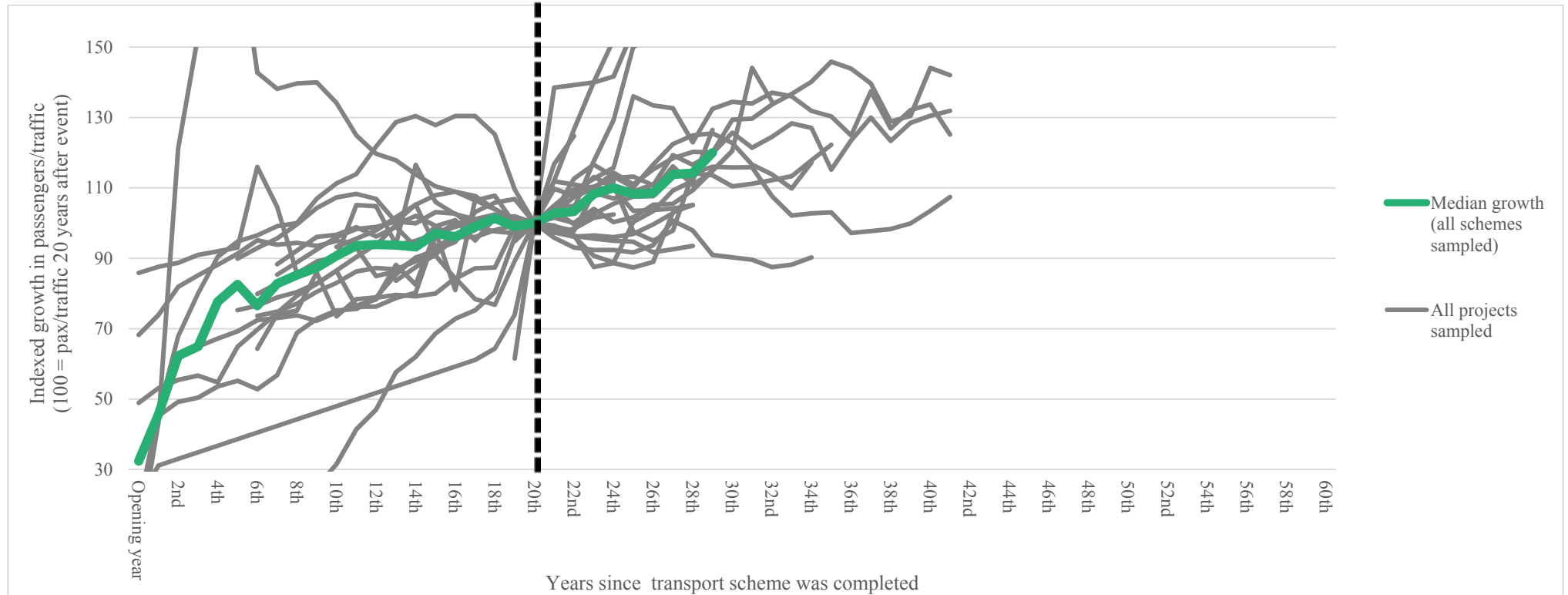


Source: DfT & Rail Industry Monitor

Figure 47 shows in green the median growth profile of all schemes sampled across road and rail. The figure demonstrates that a typical scheme tends to continue growing between their 29th and 35th year, this is in contrast with zero growth (demand cap) assumption currently recommended by WebTAG.

Beyond the 35th year we have insufficient data to show clear results, but the limited data points we do have pointed to varied trajectory's, but with the median growth rate being close to 0%.

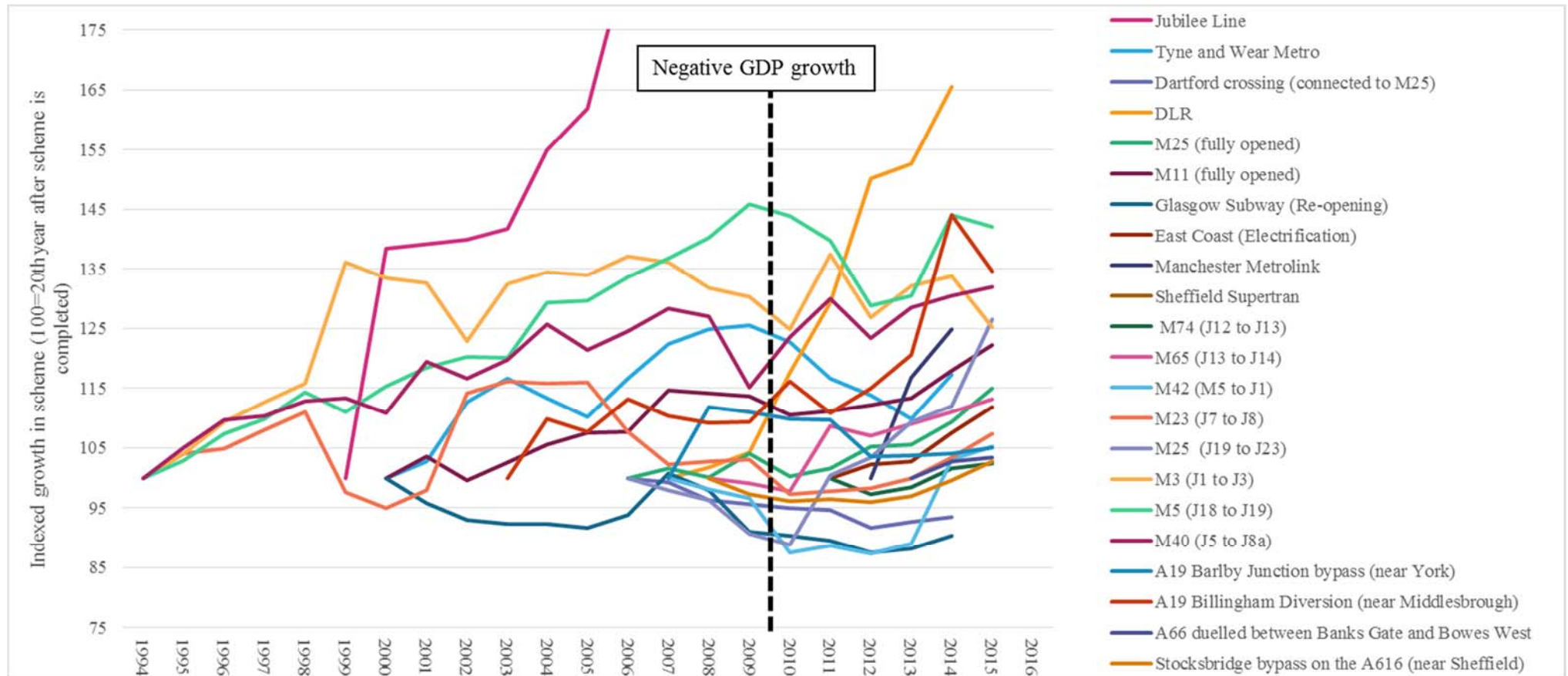
Figure 47 – Median growth profile for all schemes sampled



Source: DfT & Rail Industry Monitor

Figure 48 shows the same data as previous figures but with calendar years on the X-axis, rather than the age of the scheme.

Figure 48 – Indexed growth in patronage 20 years after transport schemes was completed



Source: DfT & Rail Industry Monitor

Individual Demand Charts

This section examines each scheme individually, establishing how demand and service provision have changed over time.

Table 17 – Full list of schemes examined as part of this project

Scheme	Year open	Mode	Included in summary analysis / charts	Reason for not being included in summary analysis / charts
A19 Barlby Junction bypass (near York)	1987	Road	Yes	n/a
A19 Billingham Diversion (near Middlesbrough)	1983	Road	Yes	n/a
Dualling of A66 between Banks Gate and Bowes West	1993	Road	Yes	n/a
Blackpool Tram	1885	Rail	No	Insufficient data about 20th-60th year since opening
Dartford crossing (fully connected to M25)	1986	Road	Yes	n/a
Dartford crossing (west tunnel opens)	1963	Road	Yes	n/a
Development of rail around Canary Wharf	Various	Rail	No	A collection of schemes - not a single scheme
Development of rail around Stratford	Various	Rail	No	A collection of schemes - not a single scheme
Docklands Light Railway (DLR);	1987	Rail	Yes	n/a
InterCity East Coast (Electrification)	1991	Rail	Yes	n/a
Eurostar	1995	Rail	No	Has not been open for more than 20 years
Glasgow Subway (Re-opening)	1980	Rail	Yes	n/a
Jubilee Line (London Underground)	1979	Rail	Yes	n/a
M1 Motorway (fully opened)	1999	Road	No	Has not been open for more than 20 years
M11 Motorway (fully opened)	1980	Road	Yes	n/a
M23 between Hooley and Mertsam (J7-J8);	1974	Road	Yes	n/a
M25 between Micklefield to South Mimms (J19 to J23);	1986	Road	Yes	n/a
M25 Motorway (fully opened)	1986	Road	Yes	n/a
M3 Motorway (fully opened)	1995	Road	No	Has not been open for more than 20 years
M3 between Sunbury to Lightwater (J1 to J3);	1974	Road	Yes	n/a
M40 between Stokenchurch to Waterstock (J5 to J8a).	1974	Road	Yes	n/a
M42 Southern Links (M5 to M42 J1);	1987	Road	Yes	n/a

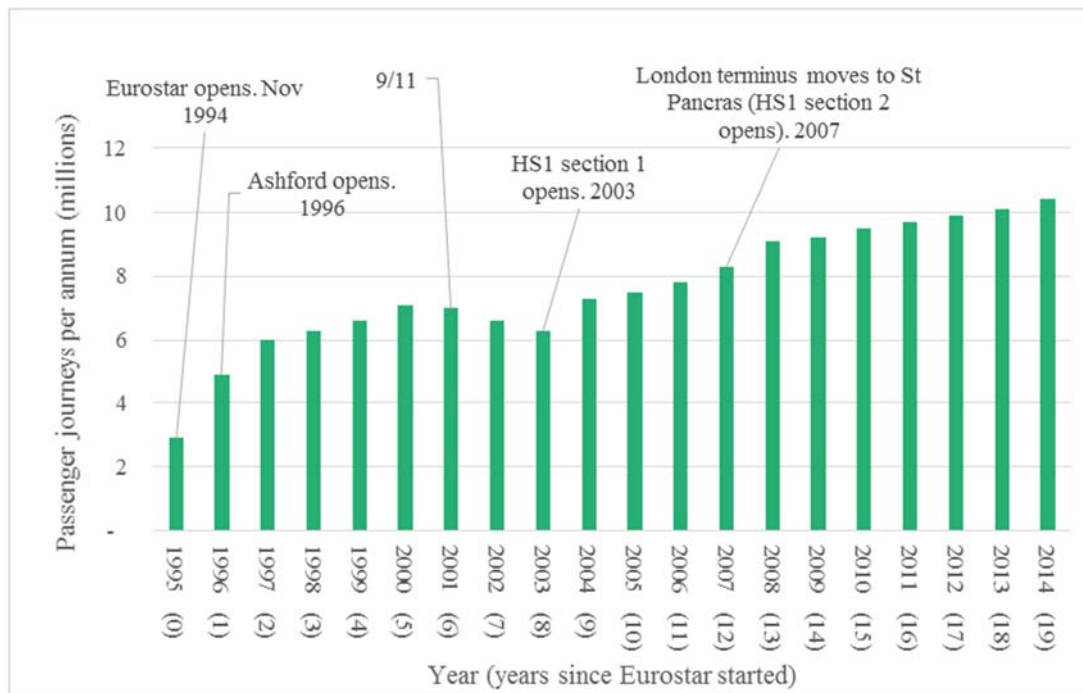
M5 between Avonmouth Bridge (J18 to J19);	1974	Road	Yes	n/a
M65 between Nelson to Colne (J13 to J14);	1988	Road	Yes	n/a
M74 between Millbank and Nether Abington (J12-J13)	1991	Road	Yes	n/a
Manchester Metrolink	1992	Rail	Yes	n/a
Nottingham Express Transit	2004	Rail	No	Has not been open for more than 20 years
Sheffield Supertran	1994	Rail	Yes	n/a
A616 Stocksbridge bypass (near Sheffield)	1988	Road	Yes	n/a
Tyne and Wear Metro	1980	Rail	Yes	n/a
Victoria Line (London Underground)	1968	Rail	No	Insufficient data about 20th-30th year since opening

Source: Arup

Eurostar

- Passenger journeys on Eurostar have grown significantly since the service was launched in 1995.
- There has been substantial infrastructure and rolling stock upgrades since launch, some of which are noted on the chart below.
- 2001 to 2003 was the only period of decline in passengers; this was predominately due to 9/11 causing a reduction in international tourism.

Figure 49 – Passenger journeys on Eurostar 1995 to 2014



Source: Booz&Co – Review of HS1 Demand Forecasts²⁰

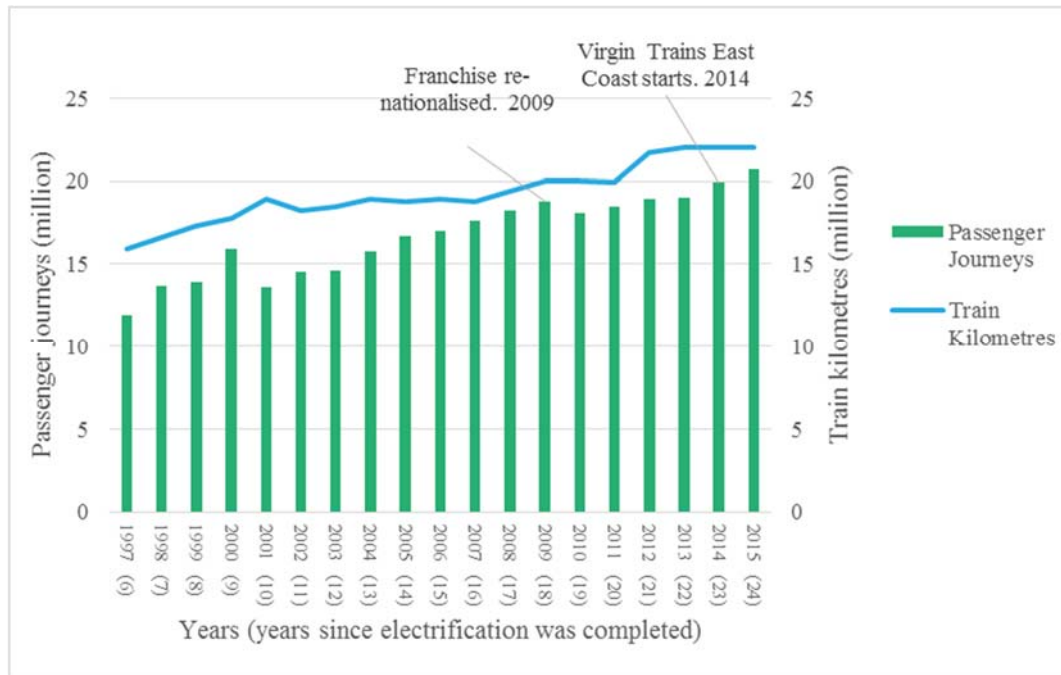
²⁰<http://assets.hs2.org.uk/sites/default/files/inserts/Review%20of%20HS1%20demand%20forecast%20s.pdf>

InterCity East Coast

East Coast Mainline opened in 1850 and was electrified between 1976 and 1991.

Passenger growth has been substantial post-electrification. Growth in passengers has broadly mirrored changes in train kilometres.

Figure 50 – Passenger journeys and train kilometres on the InterCity East Coast route, 1997 to 2015.



Source: TAS – Rail Industry Monitor²¹

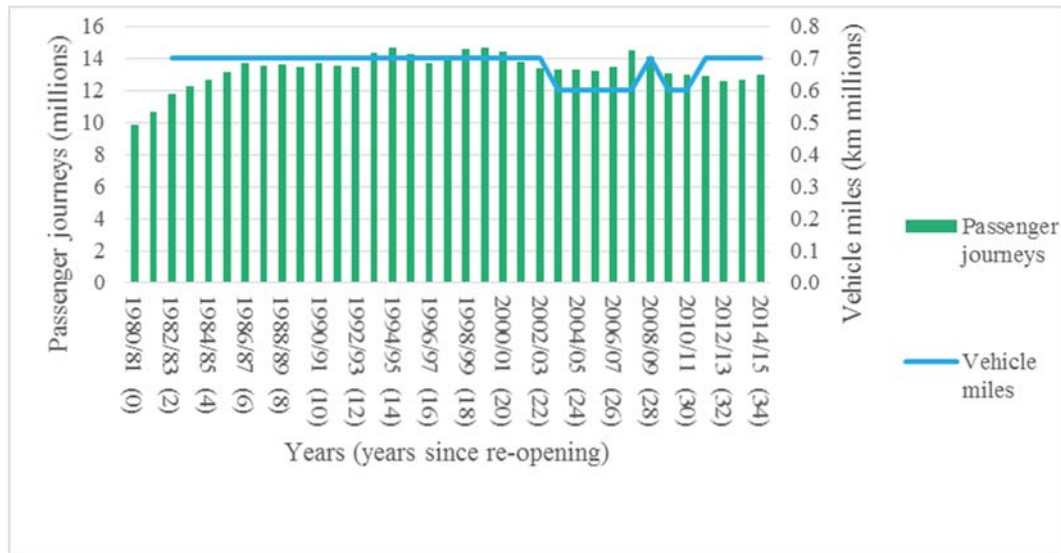
²¹ <http://www.tas-passtrans.co.uk/content/>

Glasgow subway

Glasgow Subway first opened in 1896, however during the 1970s a series of failures led to the subway being temporarily closed, re-opening in 1980.

Overall passenger numbers has been in decline over the last 10 years, with vehicle miles dropping in the 2000's, but returning to historical levels in 2010.

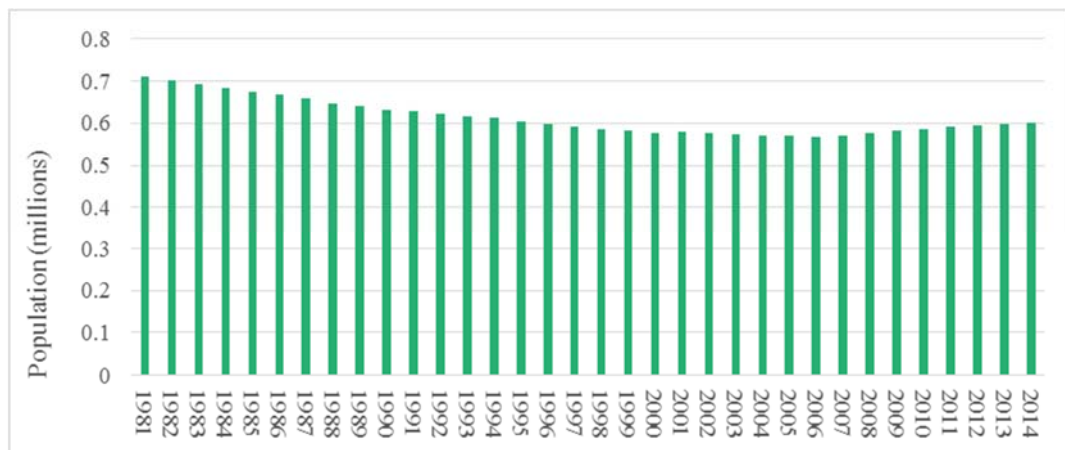
Figure 51 – Passenger journeys and vehicle miles on Glasgow Subway 1980 to 2015



(Note: Vehicle miles is only available from 1982/1983 onwards)
Source: DfT - TSGB0102²²/LRT9902a²³

The population of Glasgow City was in decline between 1980 and 2006, but has increased recently.

Figure 52 – Glasgow ²⁴ population 1981-2014



Source: ONS²⁵

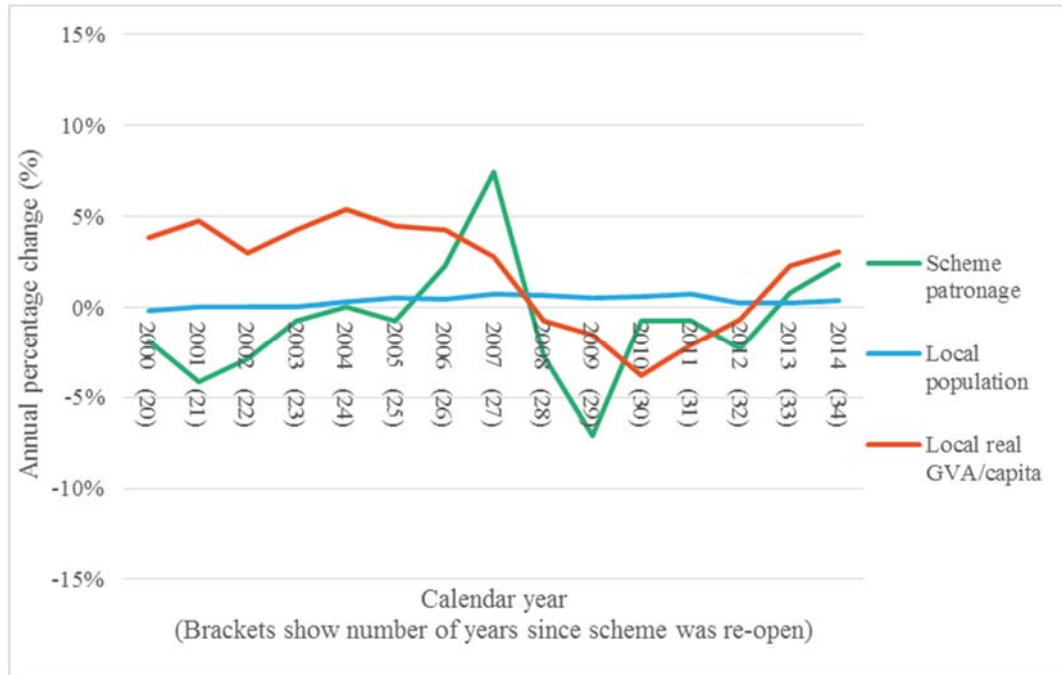
²² <https://www.gov.uk/government/statistical-data-sets/tsgb01-modal-comparisons>

²³ <https://www.gov.uk/government/statistical-data-sets/lrt99-london-glasgow-underground-train-statistics>

²⁴ LA code: S12000046

²⁵ <https://www.nomisweb.co.uk/>

Figure 53 – Growth in patronage 20 years after opening compared to growth in local GVA/capita and population - Re-opening of Glasgow Subway



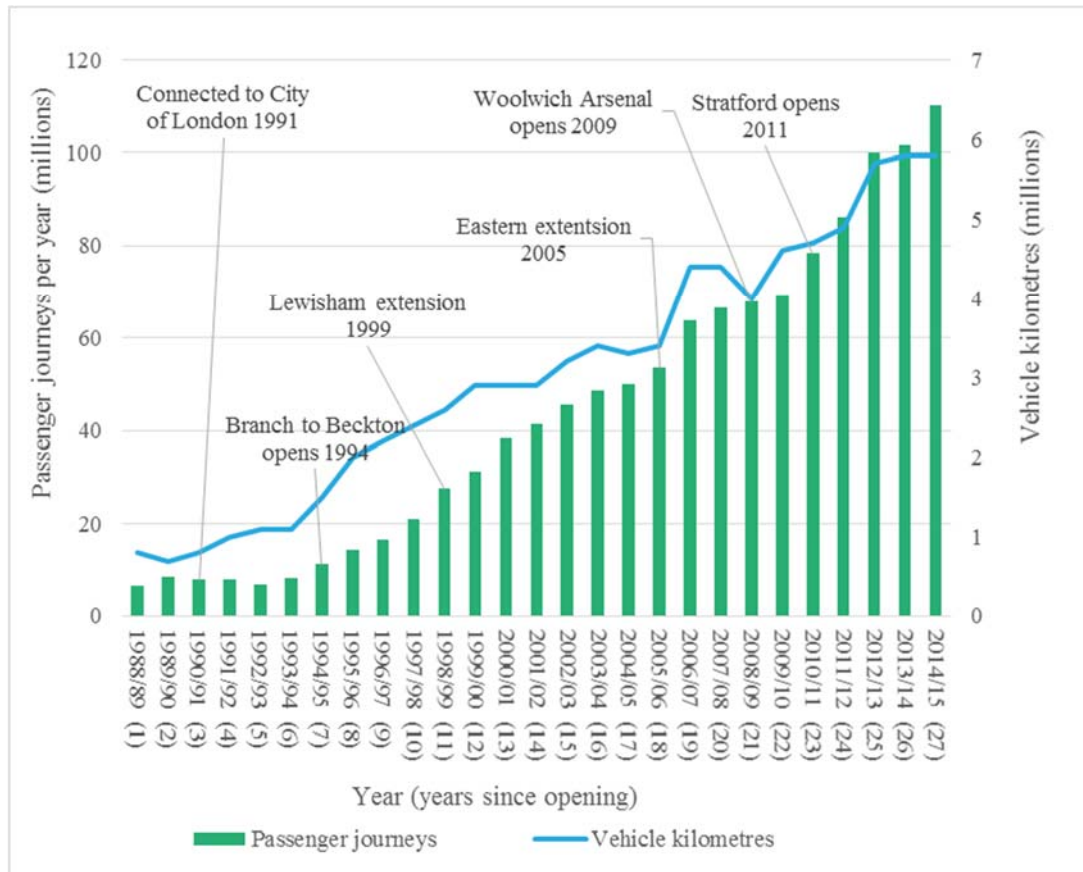
Source: DfT and ONS

DLR

Since opening in 1987 the DLR has had a number of significant upgrades and extensions, this has led to a significant increases in vehicle kilometres over time.

Passenger journeys have grown year-on-year since 1987. Twenty years after opening the service is still experiencing substantial passenger growth.

Figure 54 – Passenger journeys and vehicle kilometres on DLR 1988 to 2015



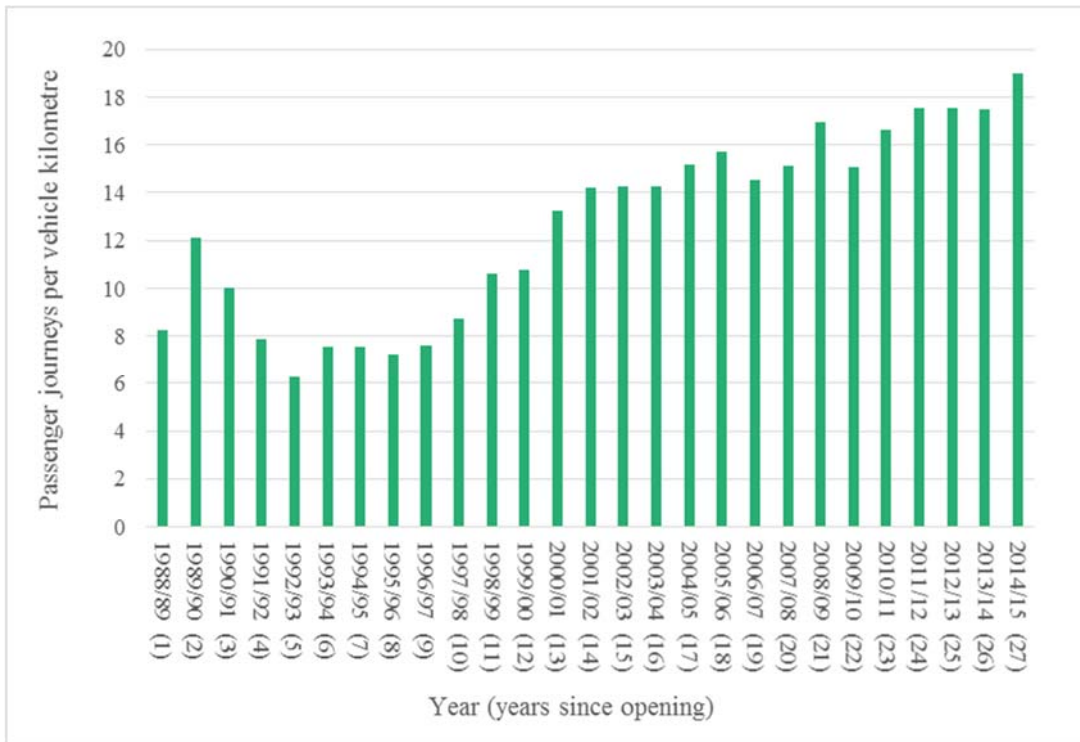
Source: DfT – Tables LRT0101²⁶/LRT0105²⁷

Part of the reason passenger growth has occurred is due to improvements in capacity and frequency – supply side changes. However, if we examine passenger journeys per vehicle kilometre travelled, we also see growth. This provides some evidence that increases in passenger journeys is not just happening because they are adding more trains and track over time.

²⁶ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

²⁷ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

Figure 55 – Passenger journeys per vehicle kilometre on the DLR, 1988 to 2015



Source: DfT – Tables LRT0101²⁸/LRT0105²⁹

²⁸ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

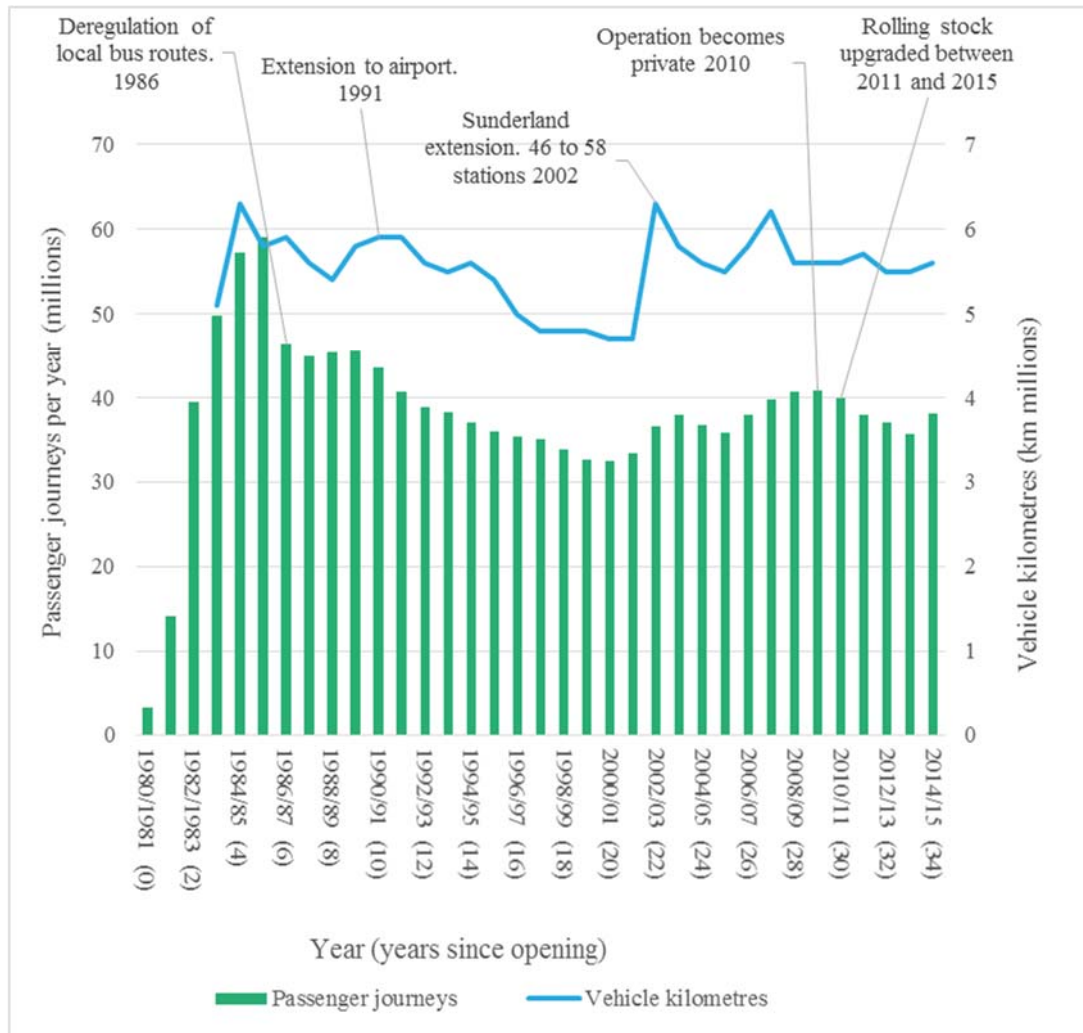
²⁹ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

Tyne and Wear Metro

Tyne and Wear Metro opened in 1980, since then it has had a number of extensions and rolling stock improvements some of which are label on the chart below.

Periods of growth and decline in passenger numbers broadly match trends in vehicle kilometres and population of the area.

Figure 56 – Passenger journeys on Tyne and Wear Metro 1980 to 2015



(Note: Vehicle kms only available from 1983 onwards)

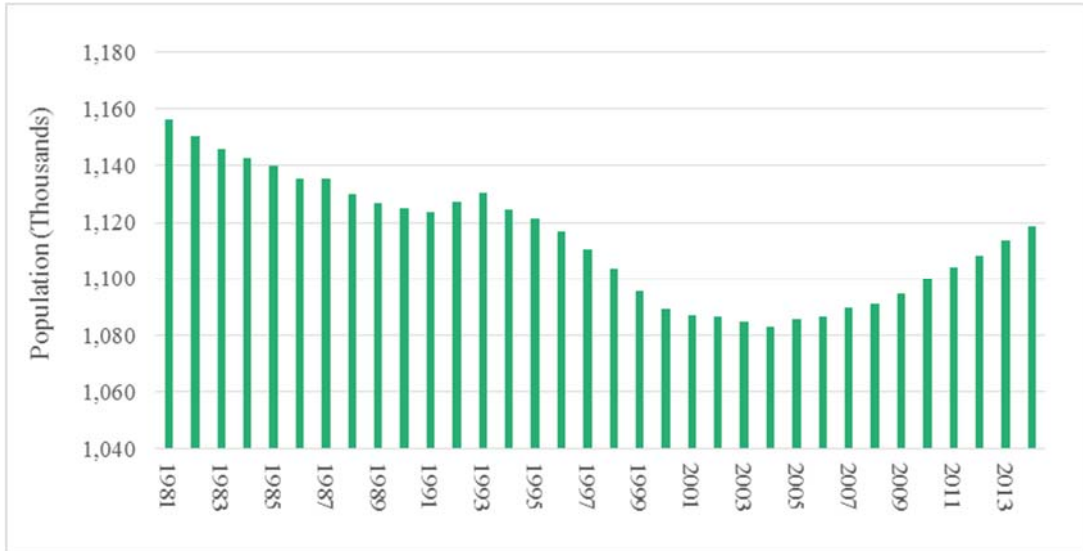
Source: DfT – Tables LRT0101³⁰/LRT0105³¹

The Tyne and Wear area had a broad decline in population until 2004 from which point the area has experienced population growth.

³⁰ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

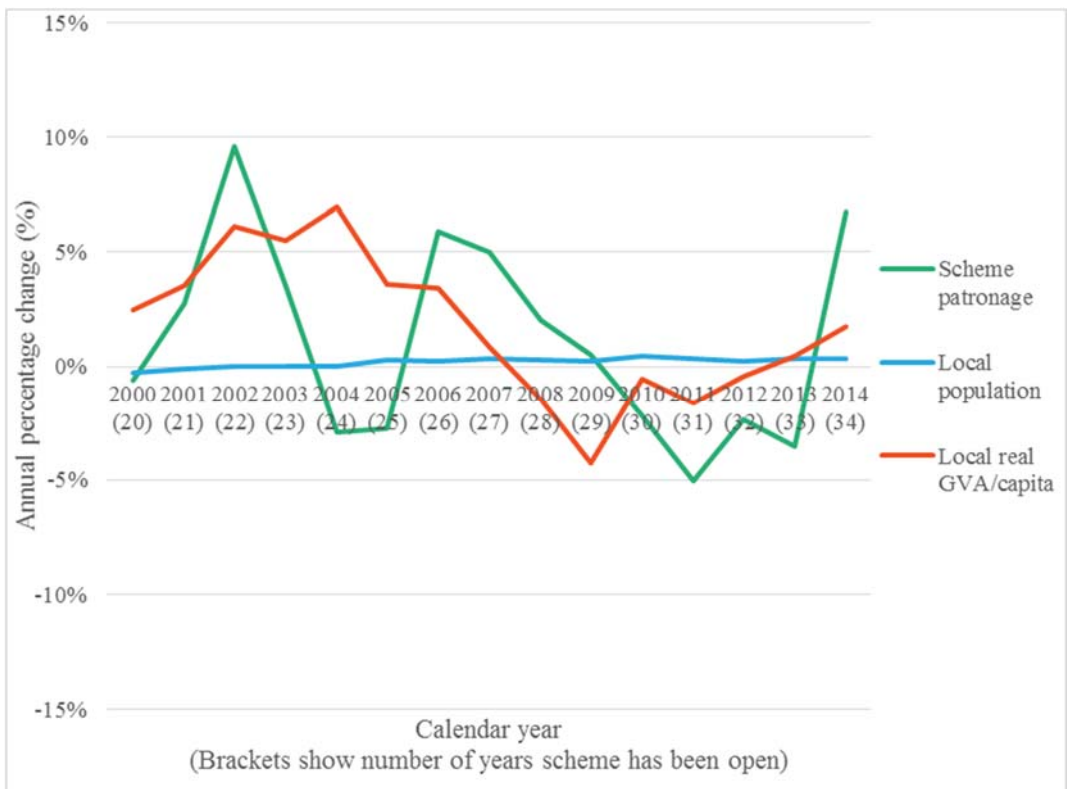
³¹ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

Figure 57 – Tyne and Wear (Met County) population 1981 to 2014



Source: ONS³²

Figure 58 – Growth in patronage 20 years after opening compared to growth in local GVA/capita and population - Tyne and Wear Metro



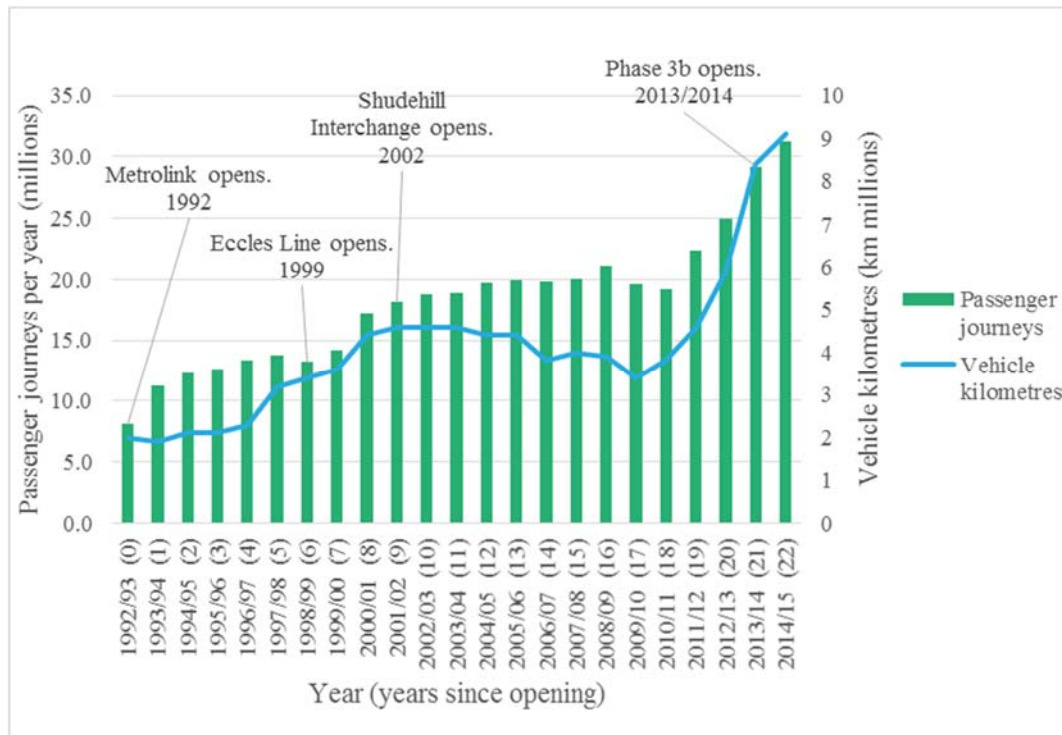
Source: DfT and ONS

³² Population is of Tyne and Wear (Met County) code: E11000004

Manchester Metrolink

Manchester Metrolink is a light rail tram system which has been open since 1992, since then it has had a number of phased extensions which have driven vehicle kilometres growth.

Figure 59 – Passenger journeys and vehicle kilometres on Manchester Metrolink 1992 to 2015



Source: DfT – Tables LRT0101³³/LRT0105³⁴

³³ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

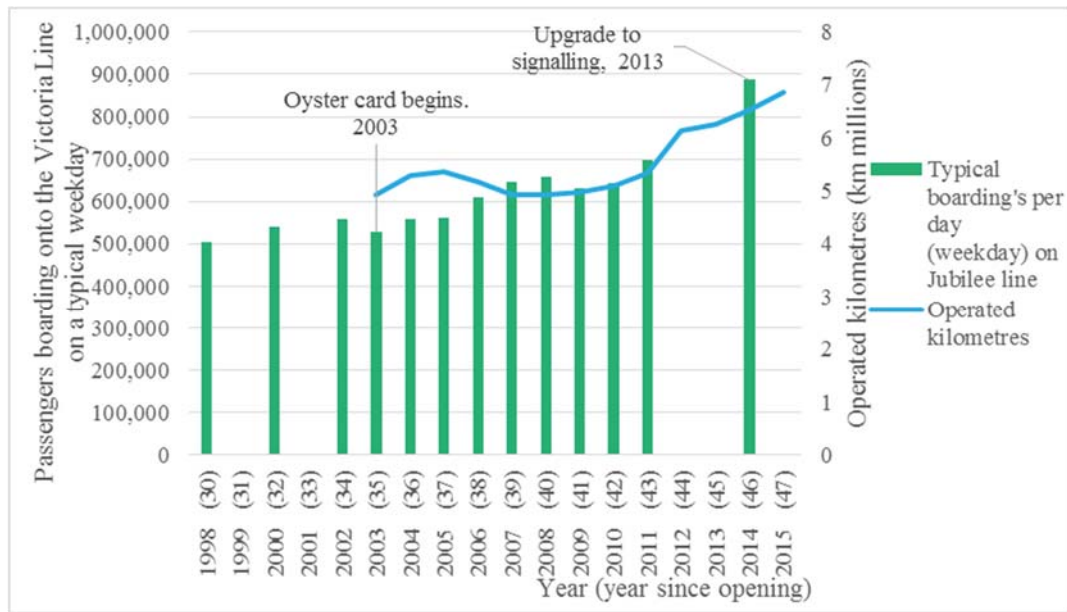
³⁴ <https://www.gov.uk/government/statistical-data-sets/lrt01-occupancy-journeys-and-passenger-miles>

Victoria Line

The Victoria Line opened in 1968, a 3.5 mile extension to Brixton was completed in 1971.

Available data suggests that usage of the Victoria Line has increased over time.

Figure 60 – Approximate number of passengers boarding onto Victoria Line services on a typical weekday and operated KMs on the line, 1998 to 2014



Source: TfL - RODS³⁵ & Performance data almanac³⁶

³⁵ https://www.whatdotheyknow.com/request/usage_of_london_underground_dlr
<https://tfl.gov.uk/info-for/open-data-users/our-feeds>

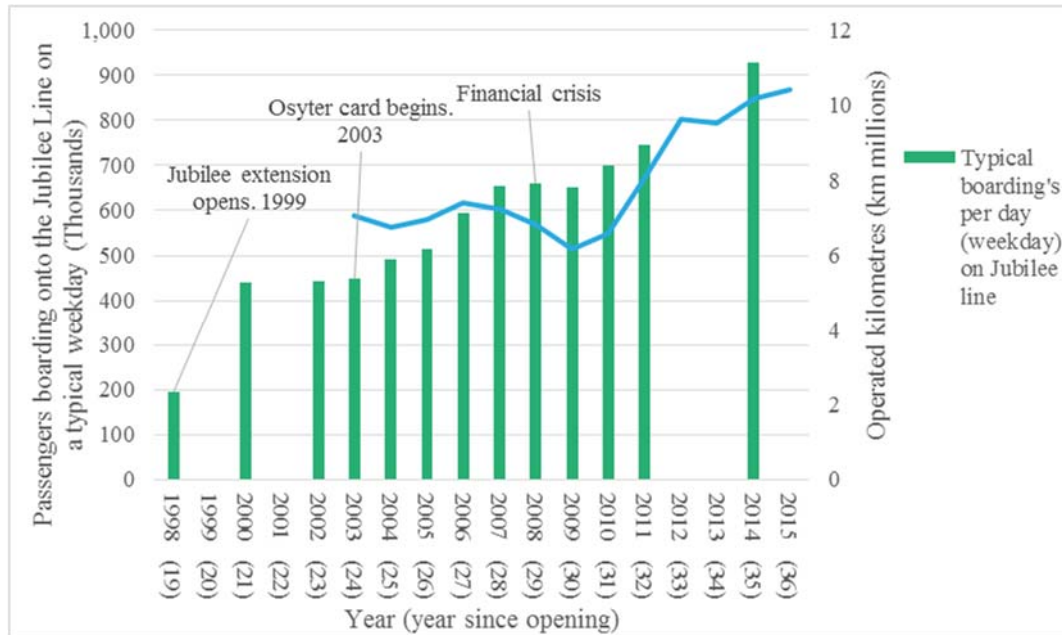
³⁶ <https://tfl.gov.uk/corporate/publications-and-reports/underground-services-performance#on-this-page-1>

Jubilee Line

The Jubilee Line was opened in 1979.

Since the extension to the Jubilee Line in 1999 the number of passenger using the service has increased significantly. A slight decline in passenger boarding's is detected during the 2008 financial crisis.

Figure 61 – Approximate number of passengers boarding onto Jubilee Line services on a typical weekday and operated KMs on the line, 1998 to 2014



Source: TfL - RODS³⁷ & Performance data almanac³⁸

³⁷ https://www.whatdotheyknow.com/request/usage_of_london_underground_dlr
<https://tfl.gov.uk/info-for/open-data-users/our-feeds>

³⁸ <https://tfl.gov.uk/corporate/publications-and-reports/underground-services-performance#on-this-page-1>

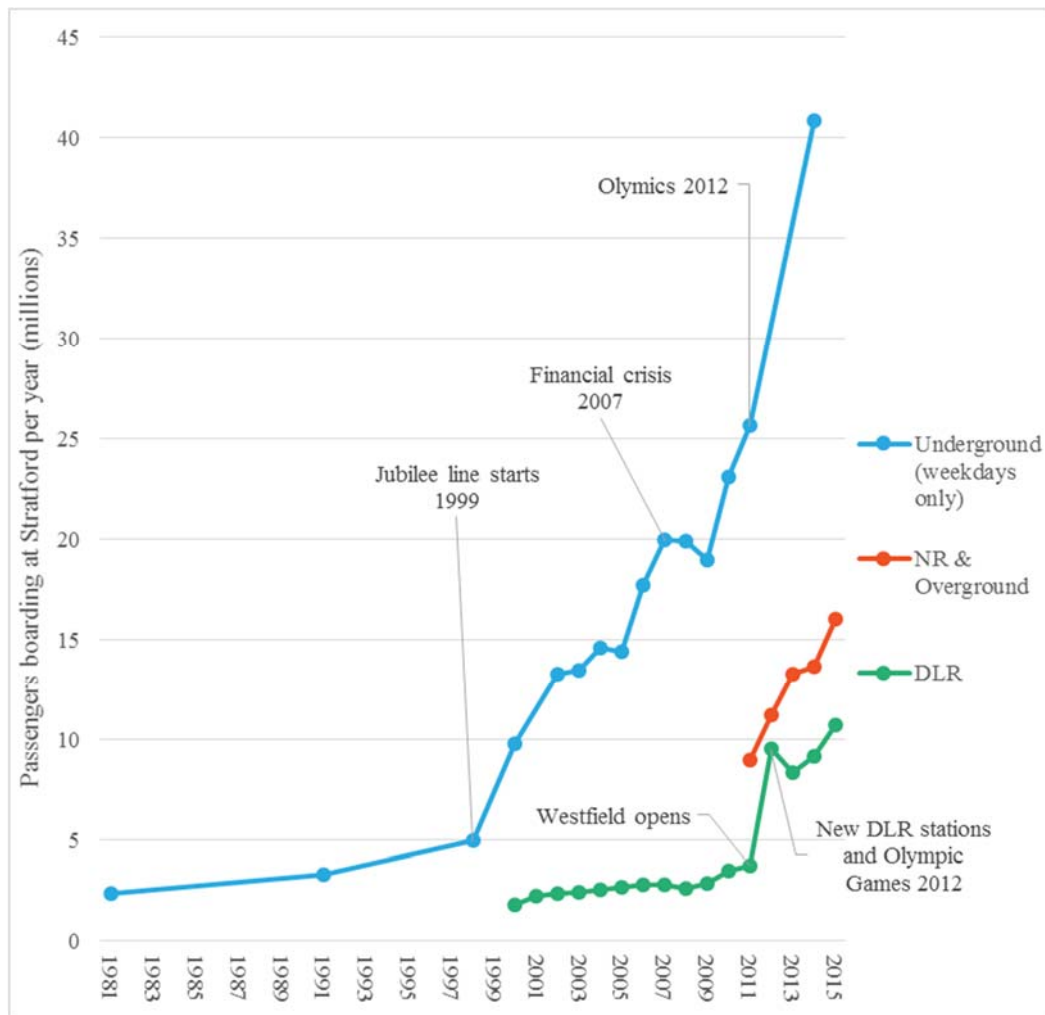
Stratford

Stratford is a significant rail interchange that includes National Rail, Overground, DLR and Underground stations.

All types of rail in Stratford have seen substantial growth over time, particularly in later years.

Passenger growth has been driven by Stratford’s increasing role as an interchange and also as a destination in itself.

Figure 62 – Approximate number of passengers boarding rail at Stratford 1981 to 2015



(Note: “Stratford” in this case study includes the following stations: Stratford, Stratford High St, Stratford International. Underground data is annualised weekday boardings only. NR & Overground data that includes Oyster card users is only available from 2010 onwards. DLR data by station is only available from 2000 onwards.)

Source: TjL³⁹ & ORR⁴⁰

³⁹ https://www.whatdotheyknow.com/request/usage_of_london_underground_dlr

⁴⁰ <http://orr.gov.uk/statistics/published-stats/station-usage-estimates>

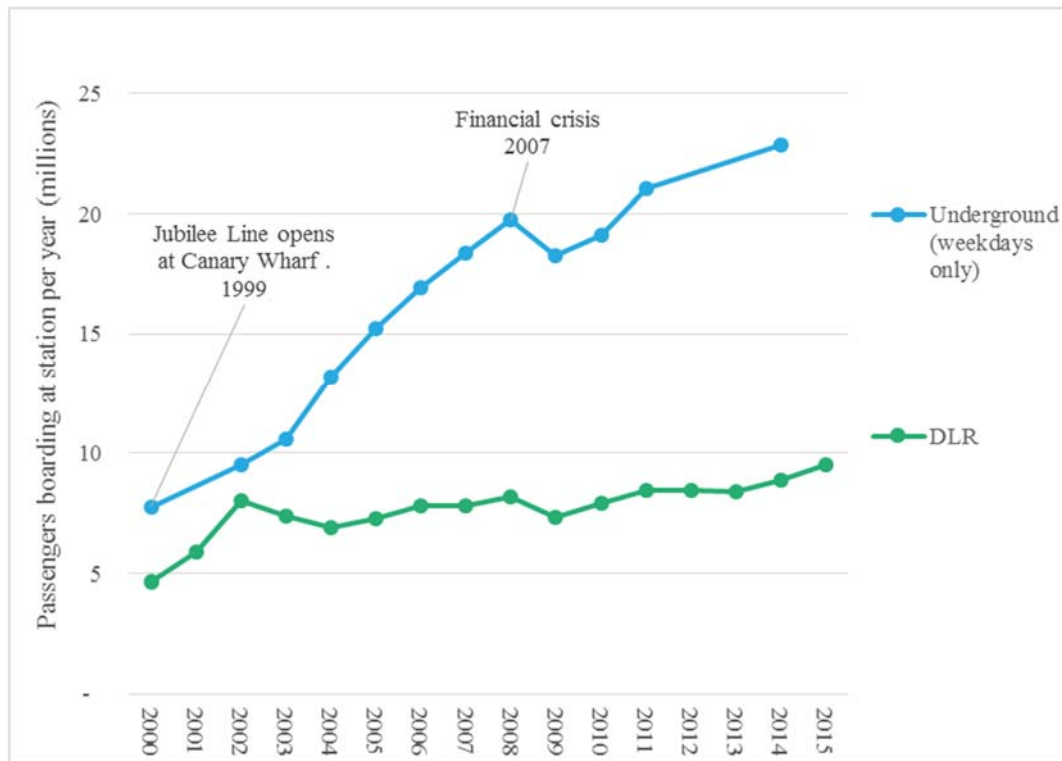
Canary Wharf

Canary Wharf DLR station opened in 1991 and Canary Wharf Underground station opened in 1999.

Construction of Canary Wharf as a development began as early as 1988.

Both have seen growth in passenger numbers, but the Underground has seen significantly greater growth.

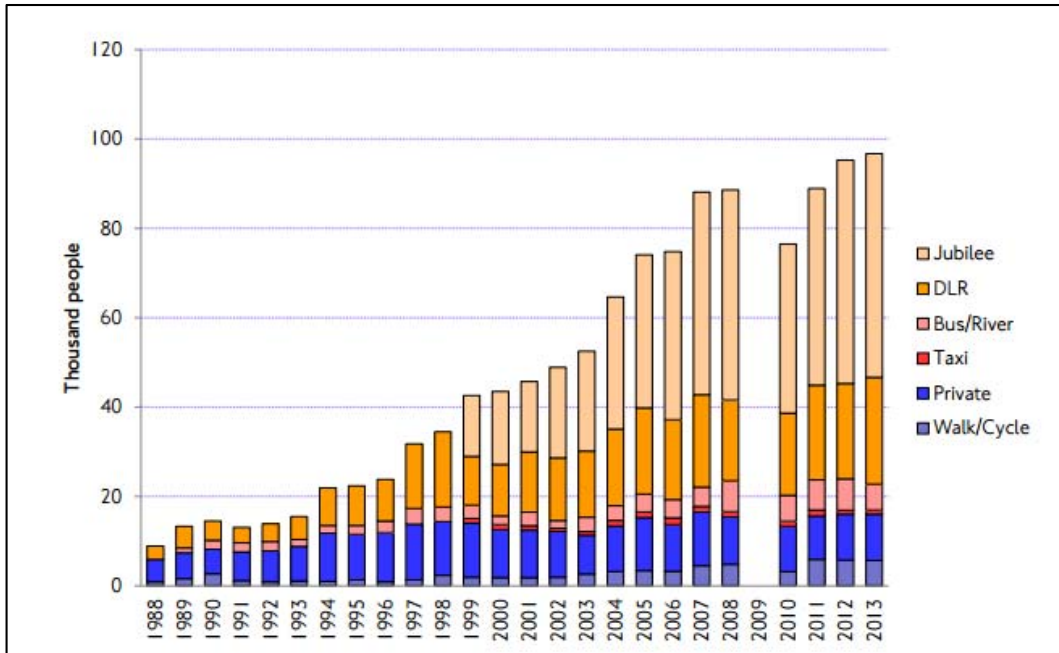
Figure 63 – Approximate number of passengers boarding rail at Canary Wharf, 2000-2015



Source: TfL⁴¹

⁴¹ https://www.whatdotheyknow.com/request/usage_of_london_underground_dlr

Figure 64 – Morning peak travel to the Isle of Dogs (including Canary Wharf) by mode of transport, 1988 to 2013



Source: TfL – Isle of Dogs Cordon Survey⁴²

- Growth in station usage has been stimulated by Canary Wharf becoming an important employment hub.

Figure 65 – Employment in Tower Hamlets 2004 to 2015



Source: ONS – Workplace analysis⁴³

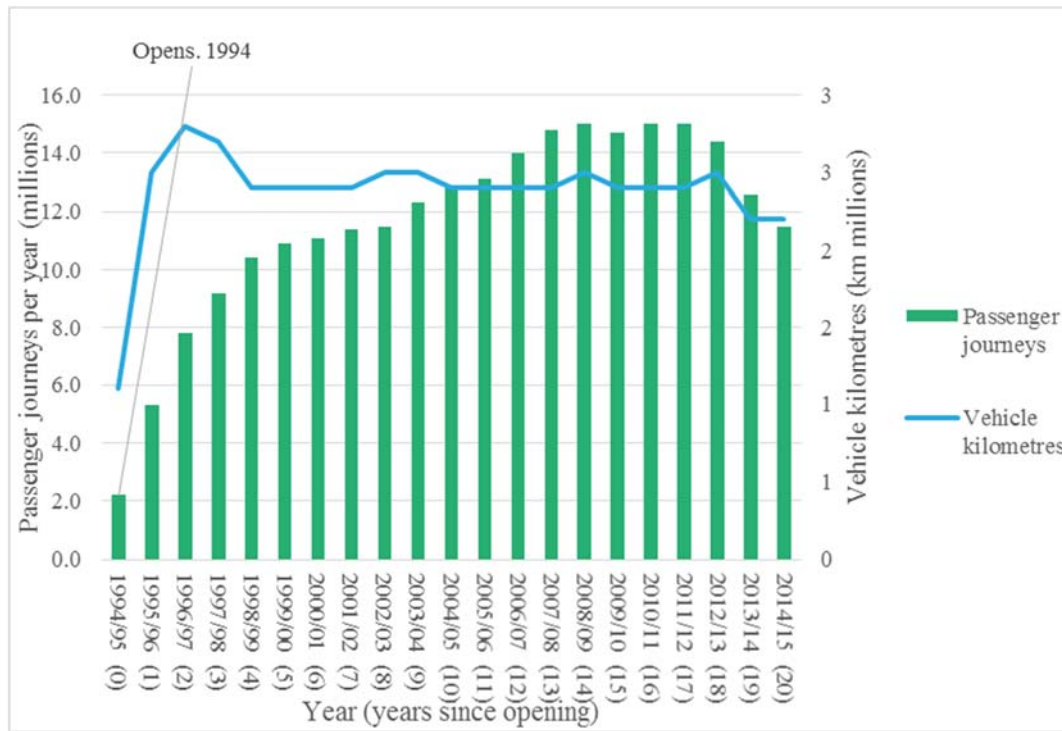
⁴² <http://content.tfl.gov.uk/isle-of-dogs-cordon-survey.pdf>

⁴³ <https://www.nomisweb.co.uk/>

Sheffield Supertram

Sheffield Supertram is a light rail tram system that was opened between 1994 and 1995. The system has not changed since opening, but a tram-train extension to Rotherham is currently under construction and is scheduled to open in 2017.

Figure 66 – Passenger journeys and vehicle kilometres on Blackpool Tram, 1963 to 2-15



Source: DfT - Table LRT0105⁴⁴/Table LRT0101⁴⁵

⁴⁴ <https://www.gov.uk/government/collections/light-rail-and-tram-statistics>

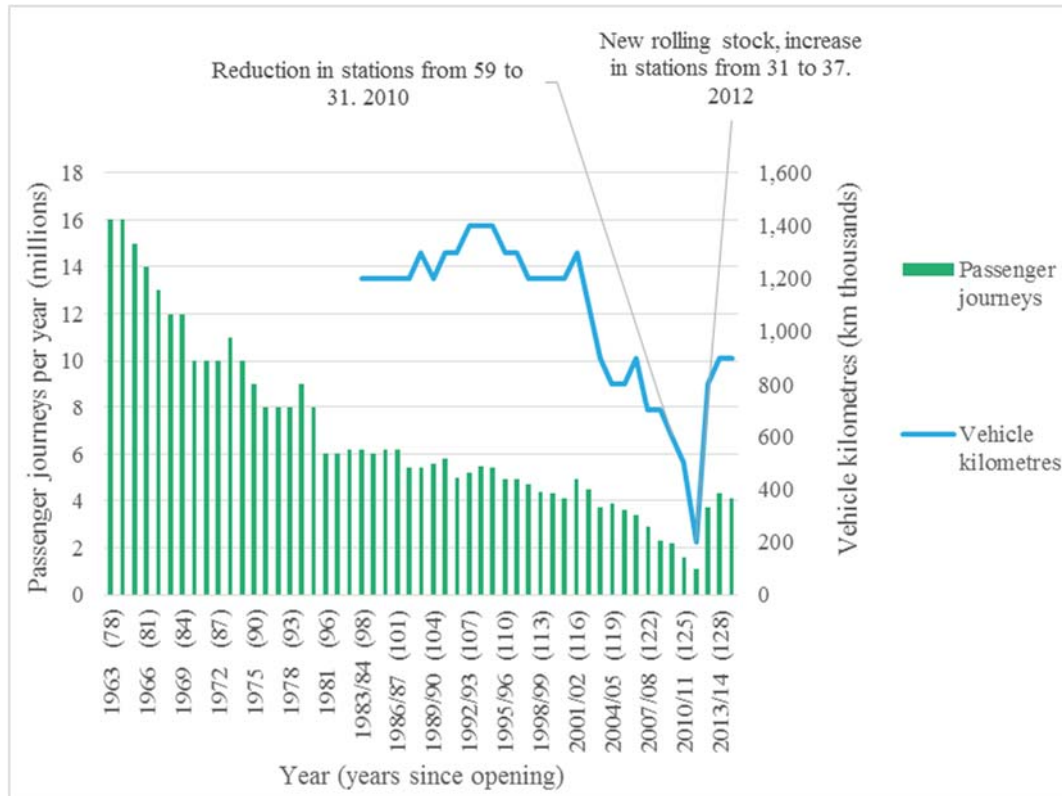
⁴⁵ <https://www.gov.uk/government/collections/light-rail-and-tram-statistics>

Blackpool Tram

Blackpool Tram first opened in 1885.

Passenger numbers have been in significant decline since the 1960s, however recent improvements to rolling stock and increases in stations have created an uplift in passenger numbers over the last couple of years.

Figure 67 – Passenger journeys and vehicle kilometres on Blackpool Tram, 1963 to 2015



(Note: Vehicle km is only available from 1983 onwards)

Source: DfT - Table LRT0105⁴⁶/Table LRT0101⁴⁷/Table TSGB0102⁴⁸

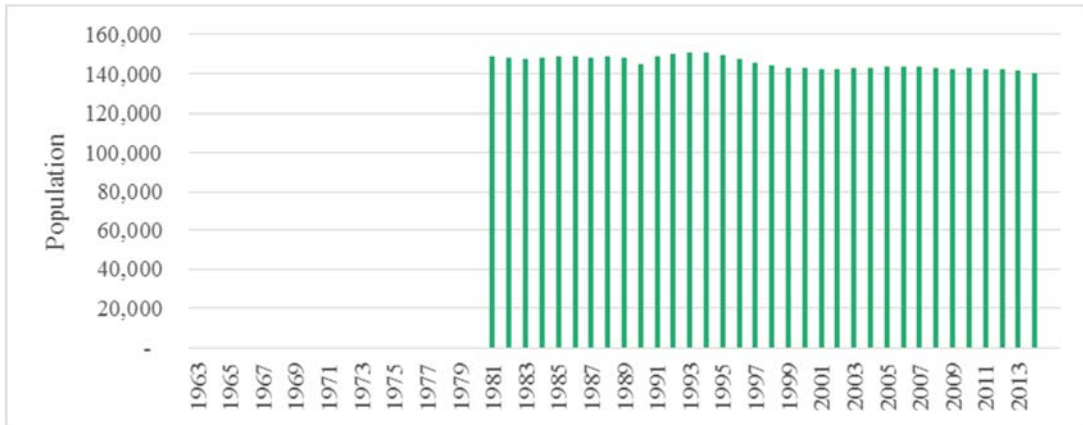
- Blackpool’s population has been decreasing slowly over time.

⁴⁶ <https://www.gov.uk/government/collections/light-rail-and-tram-statistics>

⁴⁷ <https://www.gov.uk/government/collections/light-rail-and-tram-statistics>

⁴⁸ <https://www.gov.uk/government/statistical-data-sets/tsgb01-modal-comparisons>

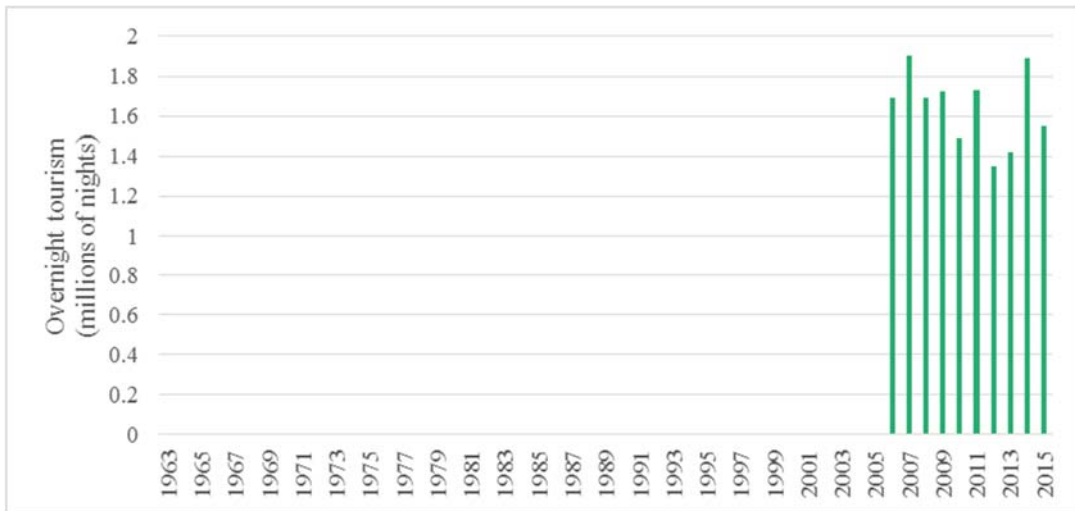
Figure 68 – Blackpool population 1981 to 2015 (Note: data before 1983 was not available)



Source: ONS

Overnight stays by people on holiday or visiting friends/relatives has been volatile but generally on a downwards trajectory.

Figure 69 – Total overnight stays in Blackpool from people on holiday or visiting friends/relatives, 2006 to 2015



(Note: Data before 1996 was not available)

Source: Visit Britain⁴⁹

⁴⁹ <https://www.visitbritain.org/nation-region-county-data>

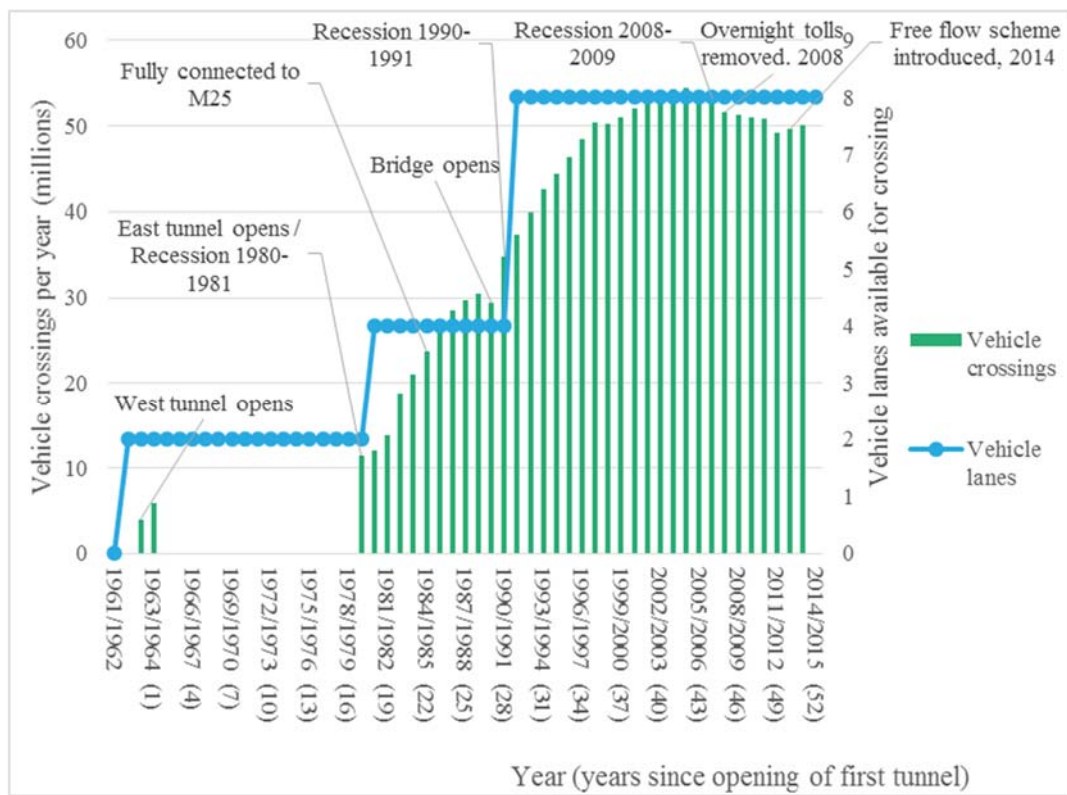
Dartford Crossing

Dartford Crossing opened in 1962 as a tunnel under the Thames. A second tunnel and then a bridge were added in 1980 and 1991 respectively.

Growth in vehicles making the crossing has increased significantly over time, this is unsurprising given the increases in car ownership over the same period and major increases in the capacity at the crossing.

Crossings peaked in 2006 and have since been declining with a slight upwards trend over the last couple of years.

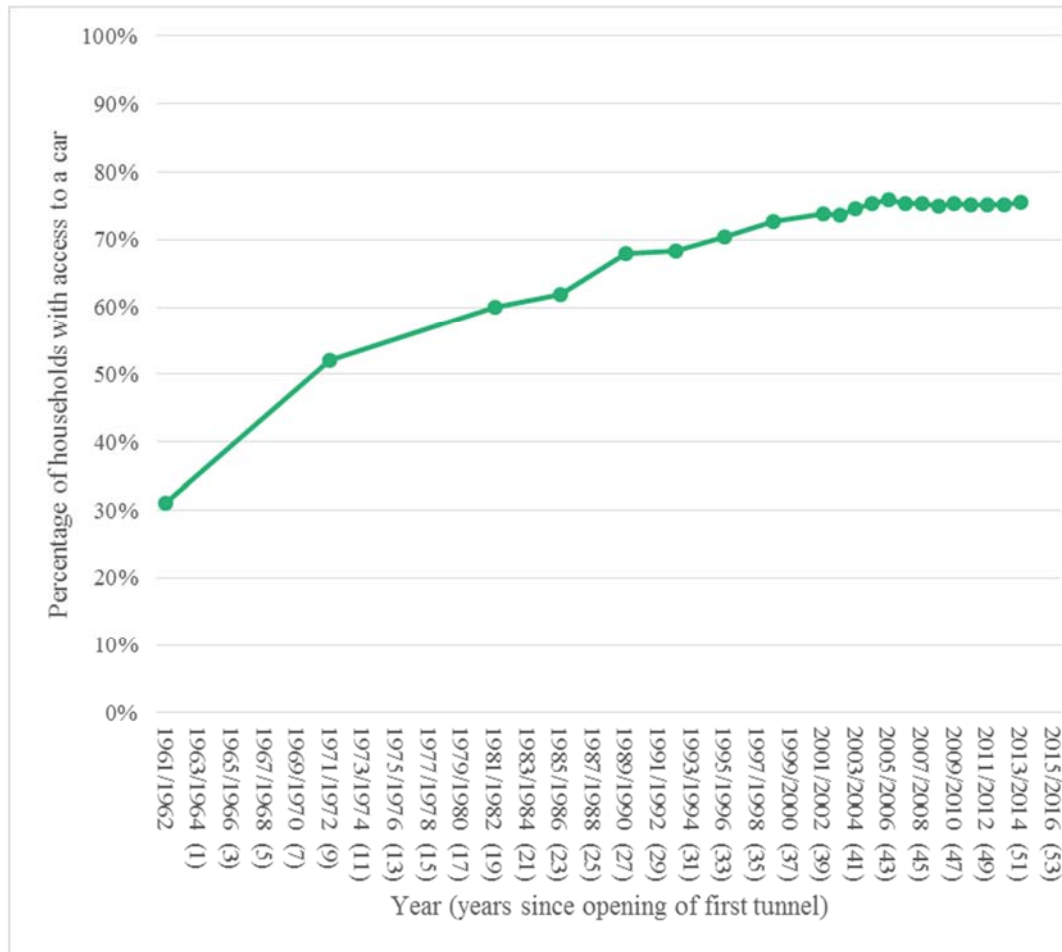
Figure 70 – Vehicles crossing at Dartford Crossing (tunnel or bridge) per annum, 1961 to 2015



Source: DfT - Traffic flows⁵⁰ (Note: Data for 1965 to 1979 was not available)

⁵⁰https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/445125/Traffic_Flow.pdf

Figure 71 – Percentage of households in England with access to a car 1961 to 2015



Source: DfT – Table NTS0205⁵¹

⁵¹ <https://www.gov.uk/government/collections/national-travel-survey-statistics>

M3 Motorway

The M3 was constructed between 1971 and 1995.

Traffic has grown since 1995, but remained fairly stable from 2001 onwards.

Figure 72 – Approximate growth in traffic on the M3 and all roads, 1995 to 2015



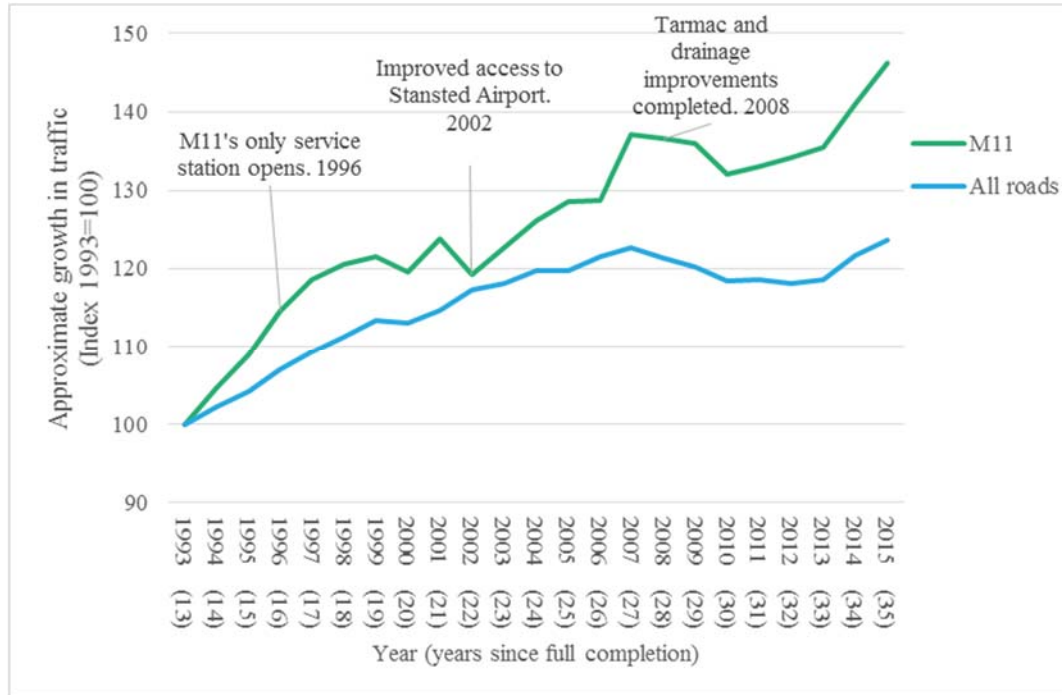
Source: Motorway data: DfT - Annual average daily flows (AADF)⁵², All roads: DfT - Table TRA0102

⁵² AADF data for 2000 to 2015 is from <https://data.gov.uk/dataset/gb-road-traffic-counts>. DfT supplied AADF data for 1993 to 1999 directly. Indexed growth shown in chart is based on aggregating growth for all DfT count points along the motorway. Only count points which have been active for the whole graphed period have been included.

M11 Motorway

- The M11 was constructed between 1975 and 1980.
- Overall traffic has steadily been increasing, with only minor improvements of that time.

Figure 73 – Approximate growth in traffic on the M11 and all roads, 1993 to 2015



Source: Motorway data: DfT - Annual average daily flows (AADF)⁵³, All roads: DfT - Table TRA0102

⁵³ AADF data for 2000 to 2015 is from <https://data.gov.uk/dataset/gb-road-traffic-counts>. DfT supplied AADF data for 1993 to 1999 directly. Indexed growth shown in chart is based on aggregating growth for all DfT count points along the motorway. Only count points which have been active for the whole graphed period have been included.

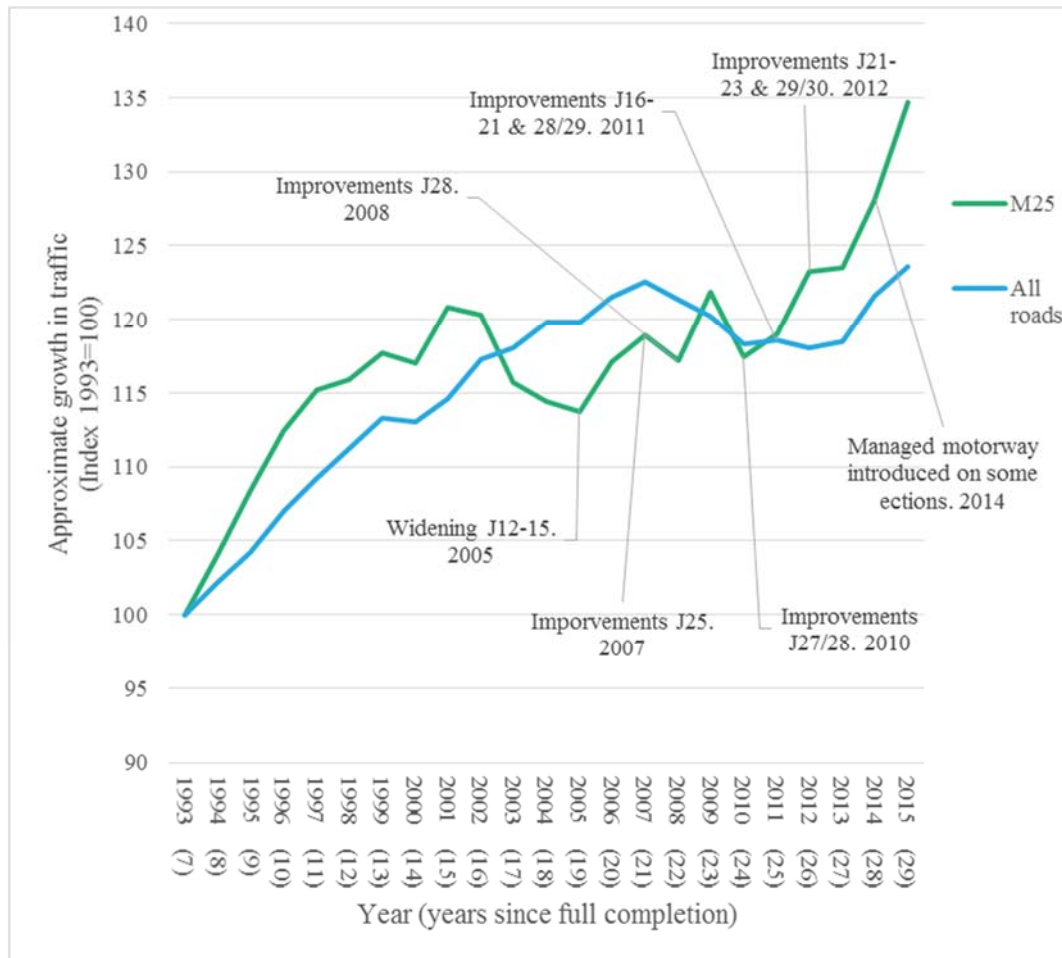
M25 Motorway

The M25 was constructed between 1975 and 1986.

Traffic grew between 1993 and 2001, remained stable between 2001 and 2010, and grew again beyond 2010.

Sections of the motorway have been improved, some of the key improvements are marked on the graph.

Figure 74 – Approximate growth in traffic on M25 and all roads, 1993 to 2015



Source: Motorway data: DfT - Annual average daily flows (AADF)⁵⁴, All roads: DfT - Table TRA0102

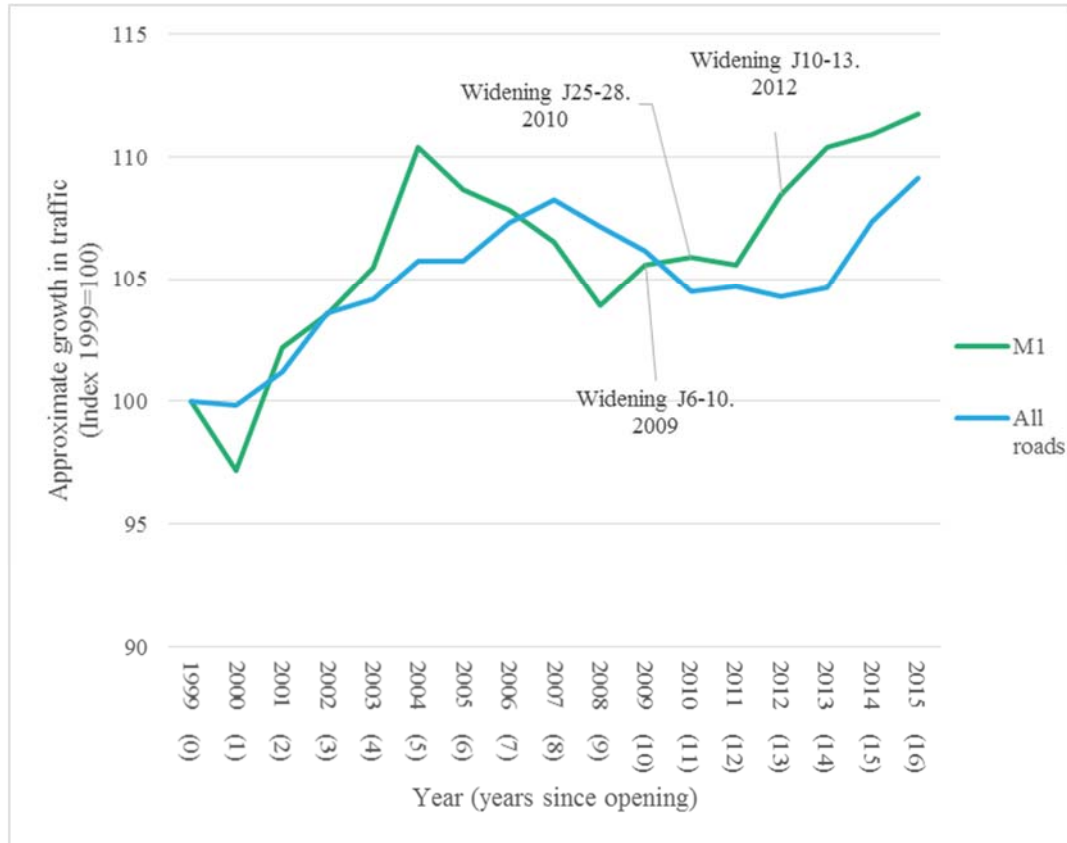
⁵⁴ AADF data for 2000 to 2015 is from <https://data.gov.uk/dataset/gb-road-traffic-counts>. DfT supplied AADF data for 1993 to 1999 directly. Indexed growth shown in chart is based on aggregating growth for all DfT count points along the motorway. Only count points which have been active for the whole graphed period have been included.

M1 Motorway

The M1 was constructed in four phases between 1959 and 1999.

Since 1999 traffic has broadly grown with some decreases between 2004 and 2008. Growth has been more modest than other motorways examined.

Figure 75 – Approximate growth in traffic on M11, 1999 to 2015



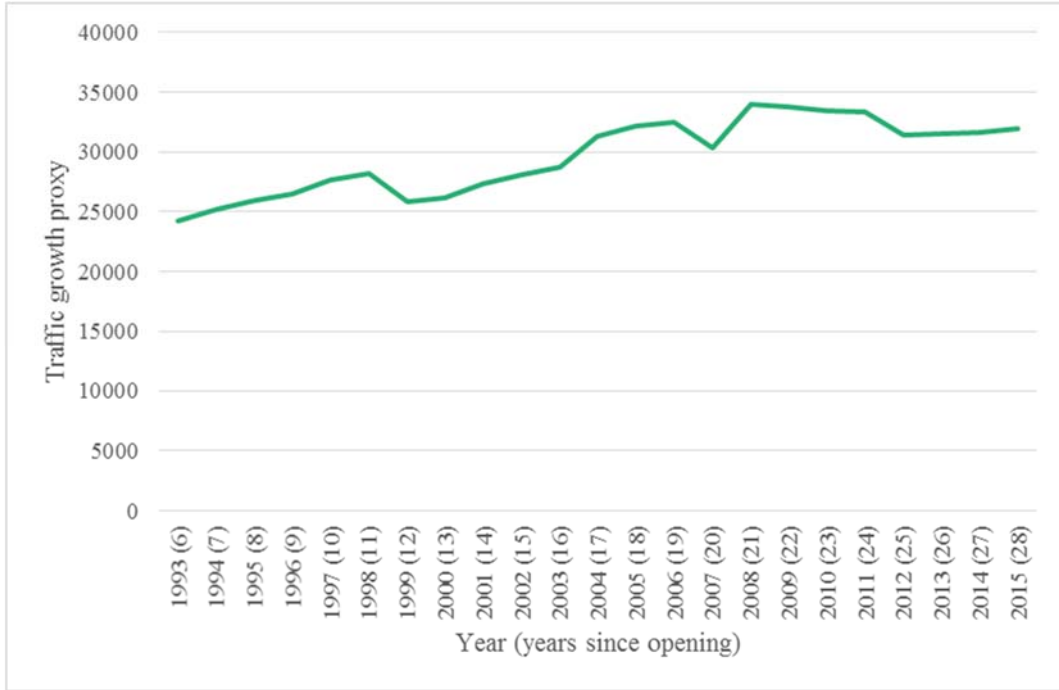
Source: Motorway data: DfT - Annual average daily flows (AADF)⁵⁵, All roads: DfT - Table TRA0102

⁵⁵ AADF data for 2000 to 2015 is from <https://data.gov.uk/dataset/gb-road-traffic-counts>. DfT supplied AADF data for 1993 to 1999 directly. Indexed growth shown in chart is based on aggregating growth for all DfT count points along the motorway. Only count points which have been active for the whole graphed period have been included.

A19 Barlby Junction bypass (near York) opens

- The five mile Barlby bypass opening in 1987

Figure 76 – Growth in at the A19 Barlby Junction bypass (near York)

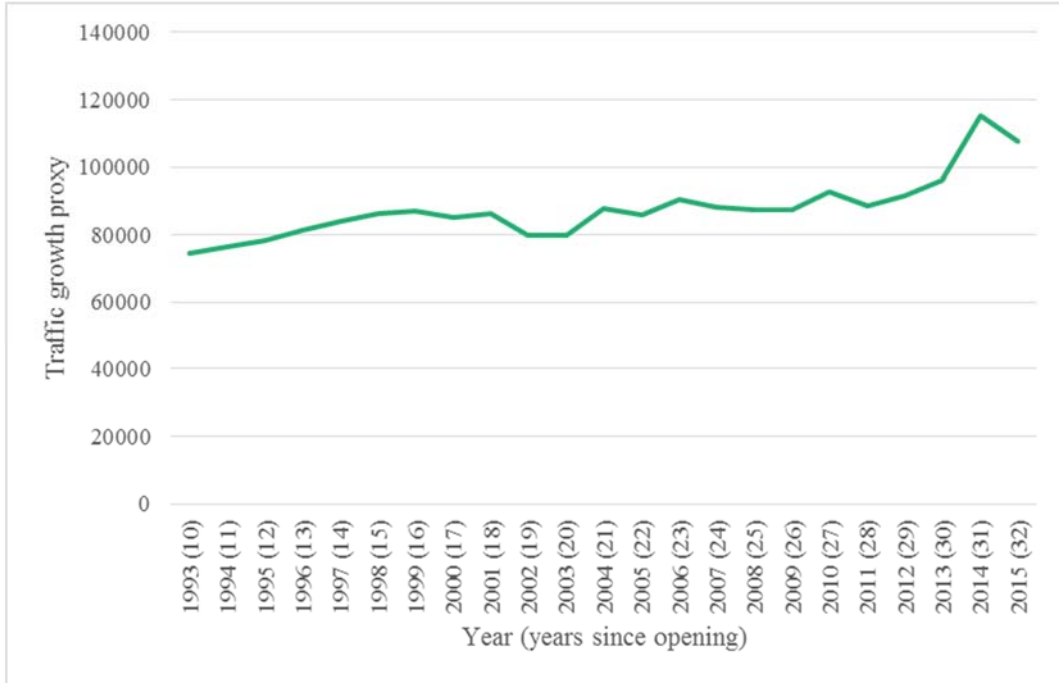


Source: DfT – AADF traffic counts

A19 Billingham Diversion (near Middlesbrough)

- The four mile Billingham Diversion was opened in 1983

Figure 77 – Growth in traffic at the A19 Billingham Diversion (near Middlesbrough)

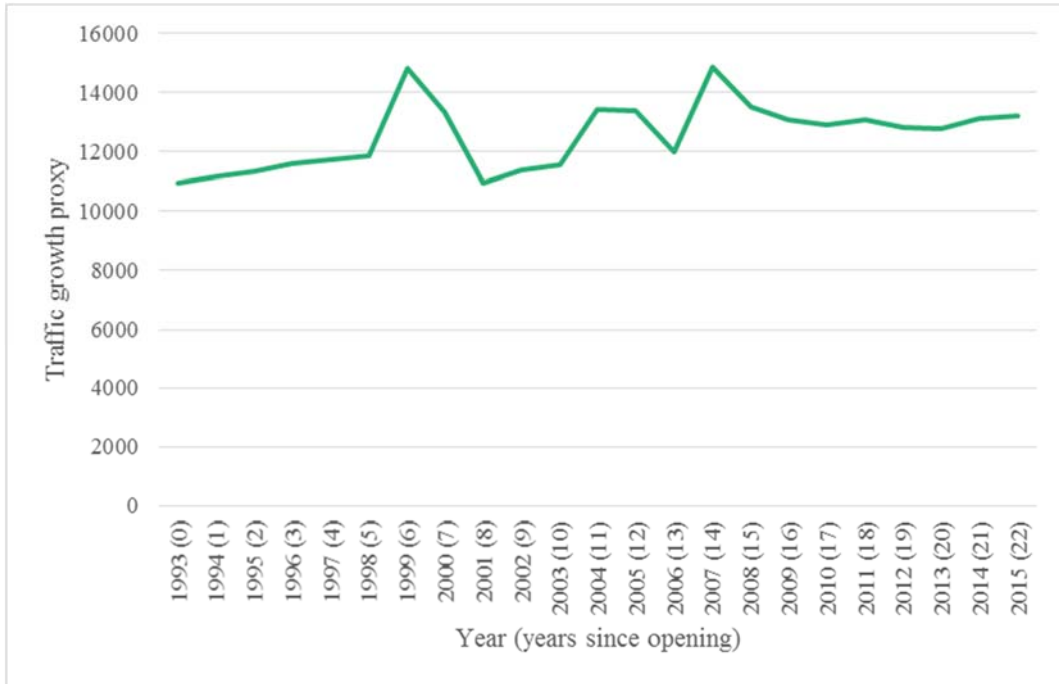


Source: DfT – AADF traffic counts

Dualling of A66 between Banks Gate and Bowes West

- The new dual-carriageway A66 between Bank Gate and Bowes West was opened in 1993.

Figure 78 – Growth in traffic after the dualling of the A66 between Banks Gate and Bowes West

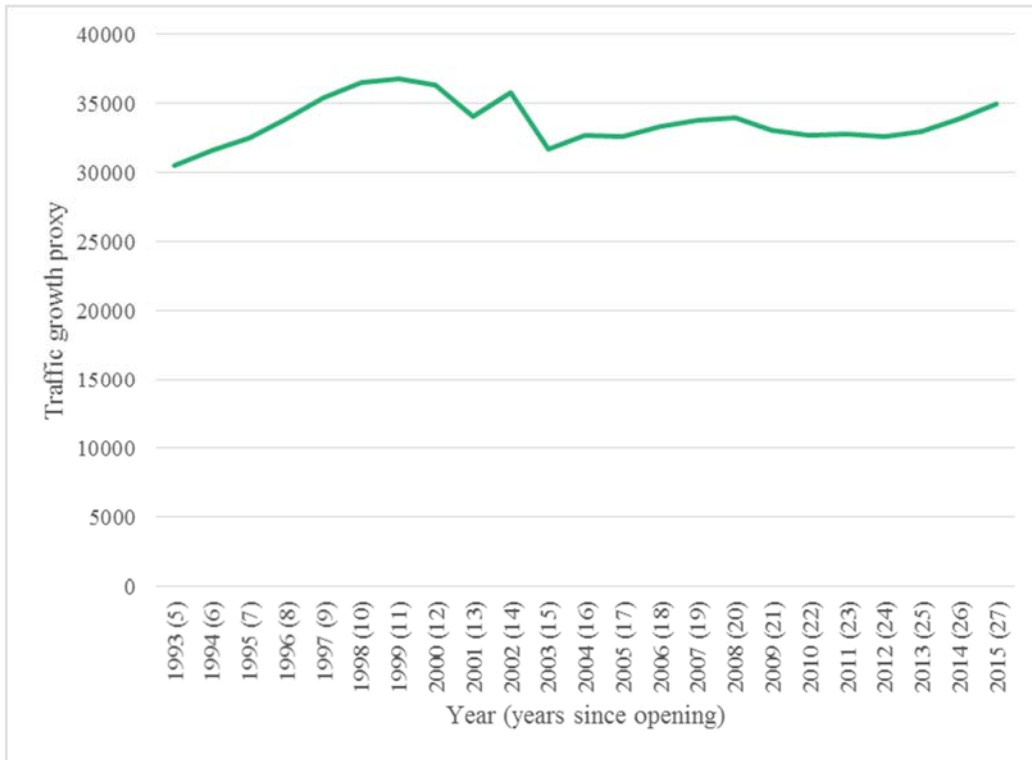


Source: DfT – AADF traffic counts

A616 Stocksbridge bypass (near Sheffield)

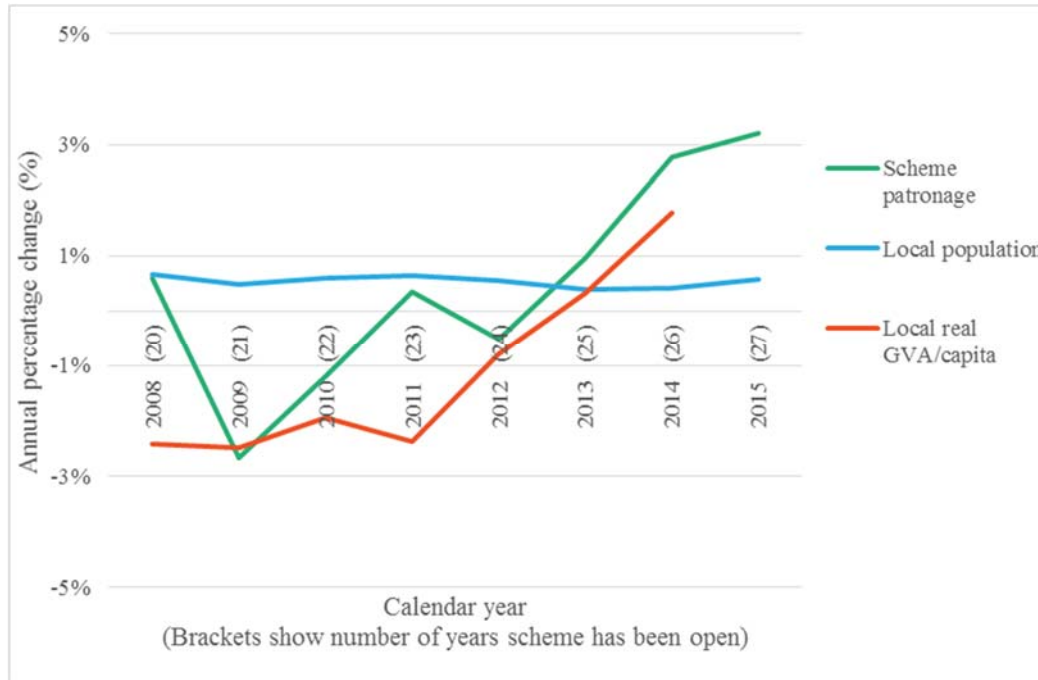
- The Stocksbridge bypass was opened in 1988

Figure 79 – Growth in traffic at the Stocksbridge bypass on the A616 (near Sheffield)



Source: DfT – AADF traffic counts

Figure 80 – Growth in patronage 20 years after opening compared to growth in local GVA/capita and population - Stocksbridge bypass on the A616 (near Sheffield)

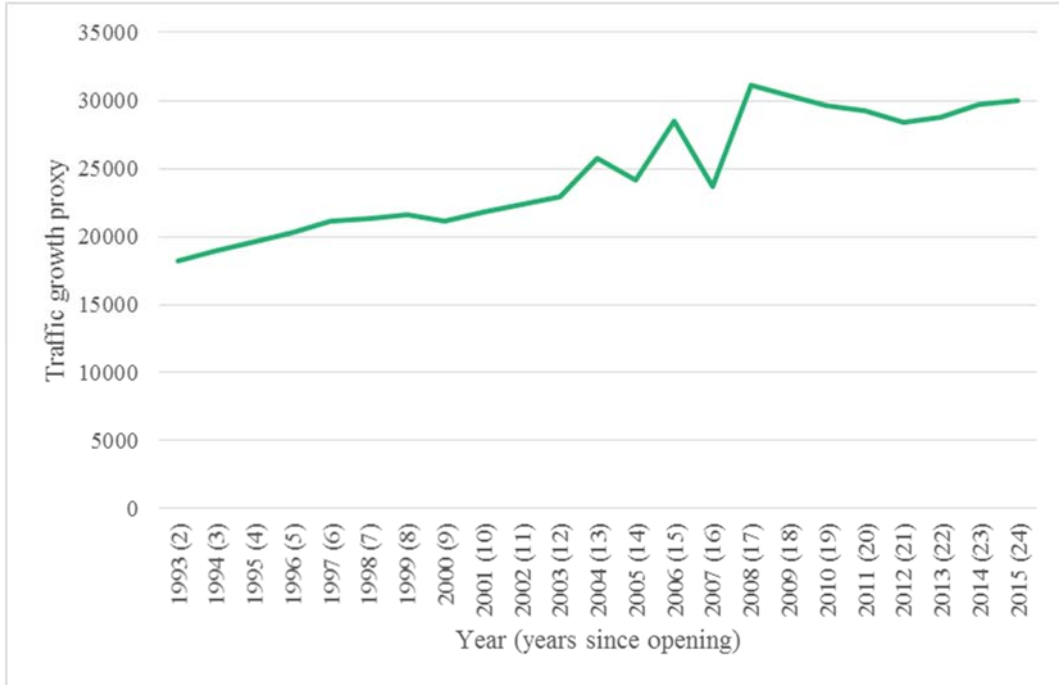


Source: DfT and ONS

M74 between Millbank and Nether Abington (J12-J13)

- This section of the M74 was opened in 1991.

Figure 81 – Growth in traffic on the M74 between J12 and J13

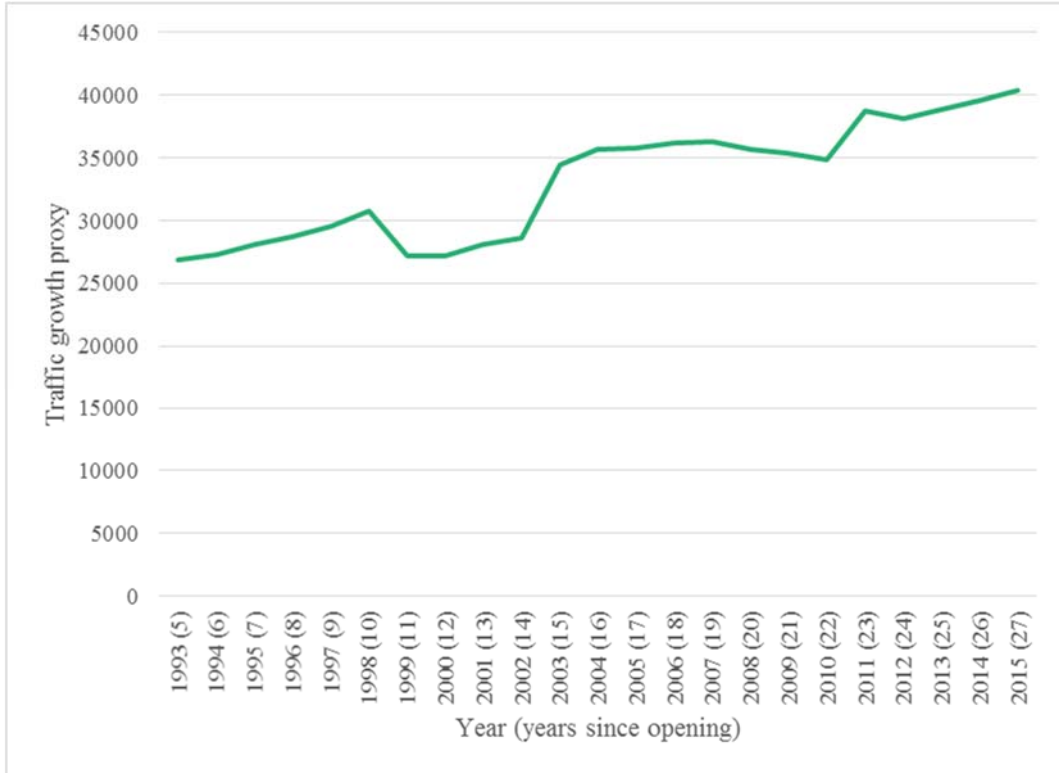


Source: DfT – AADF traffic counts

M65 between Nelson to Colne (J13 to J14)

- This section of the M65 was opened in 1988.

Figure 82 – Growth in traffic on the M65 between Nelson to Colne (J13 to J14)

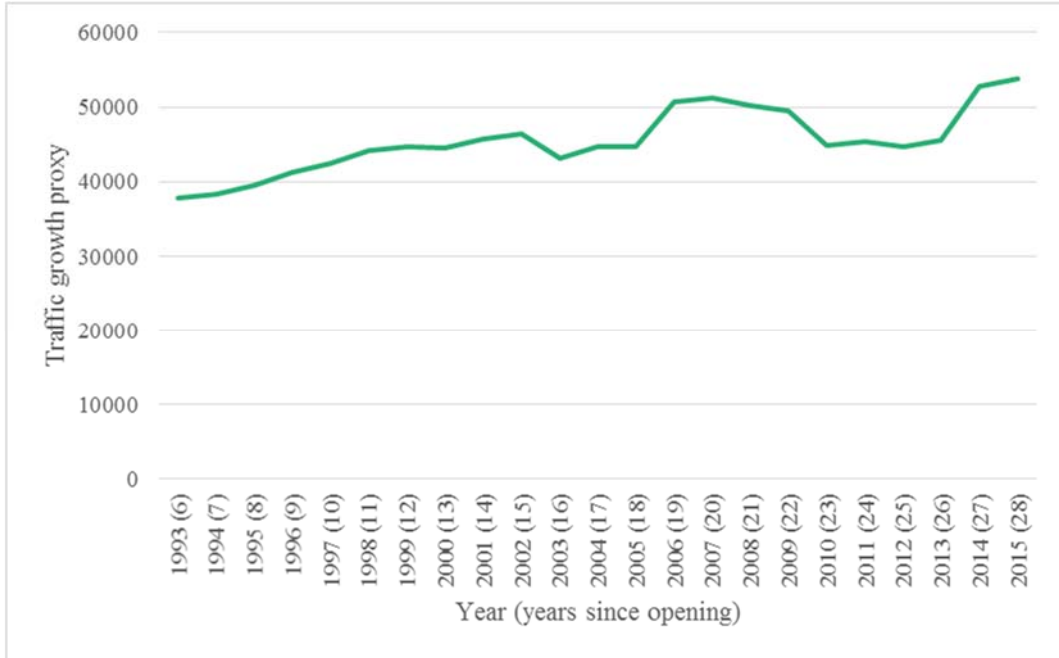


Source: DfT – AADF traffic counts

M42 Southern Links (M5 to M42 J1)

- This section of the M42 was opened in 1987.

Figure 83 – Growth in traffic at on the M42 Southern Links (M5 to M42 J1)

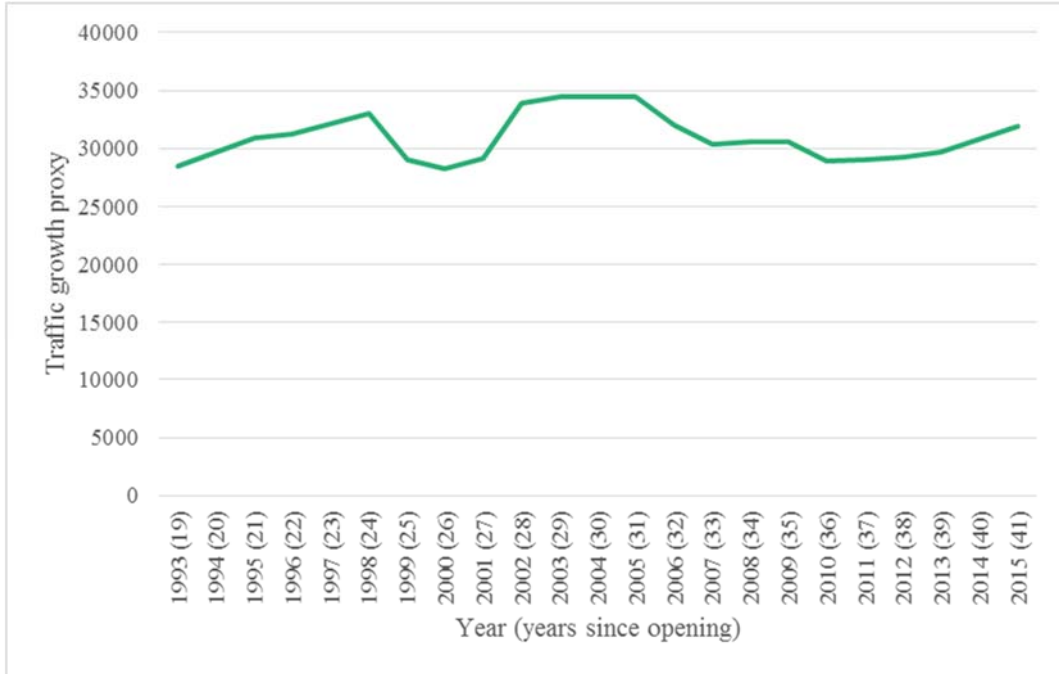


Source: DfT – AADF traffic counts

M23 between Hooley and Mertsam (J7-J8)

- This section of the M23 was opened in 1974

Figure 84 – Growth in traffic on the M23 between Hooley and Mertsam (J7-J8)

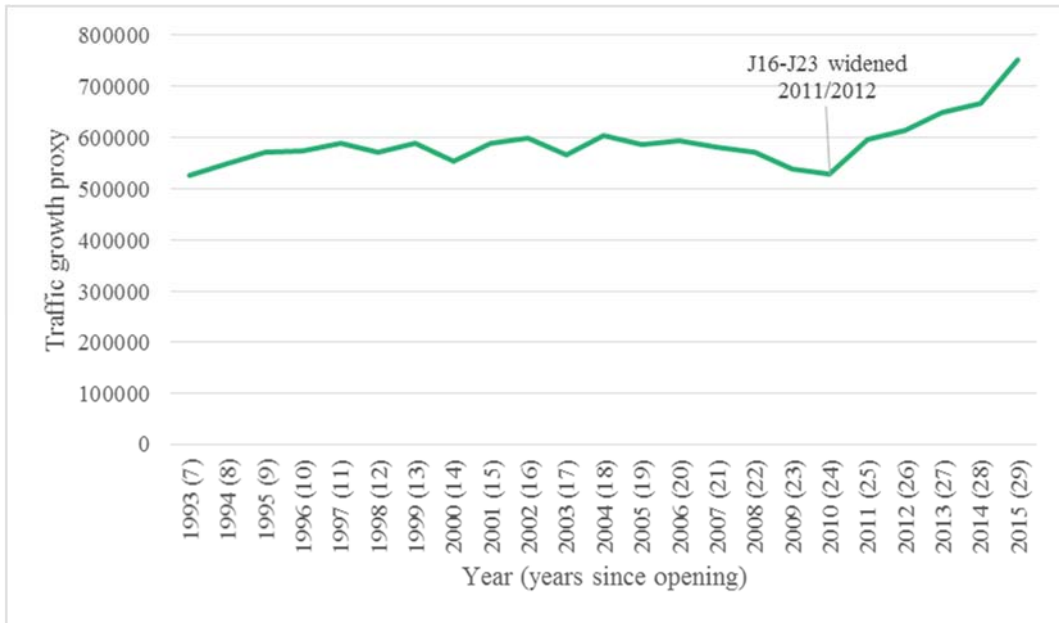


Source: DfT – AADF traffic counts

M25 between Micklefield to South Mimms (J19 to J23)

- This section of the M25 was opened in 1986

Figure 85 – Growth in traffic on the M25 between Micklefield to South Mimms (J19 to J23)

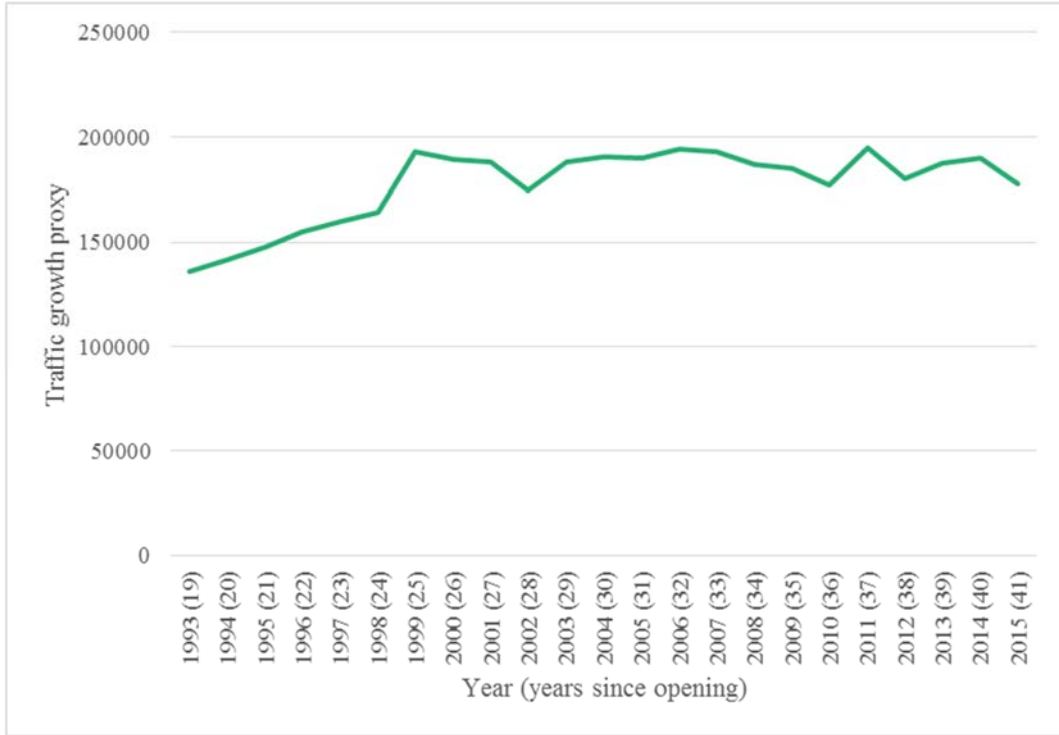


Source: DfT – AADF traffic counts

M3 between Sunbury to Lightwater (J1 to J3)

- This section of the M3 was opened in 1974

Figure 86 – Growth in traffic on the M3 between Sunbury to Lightwater (J1 to J3)

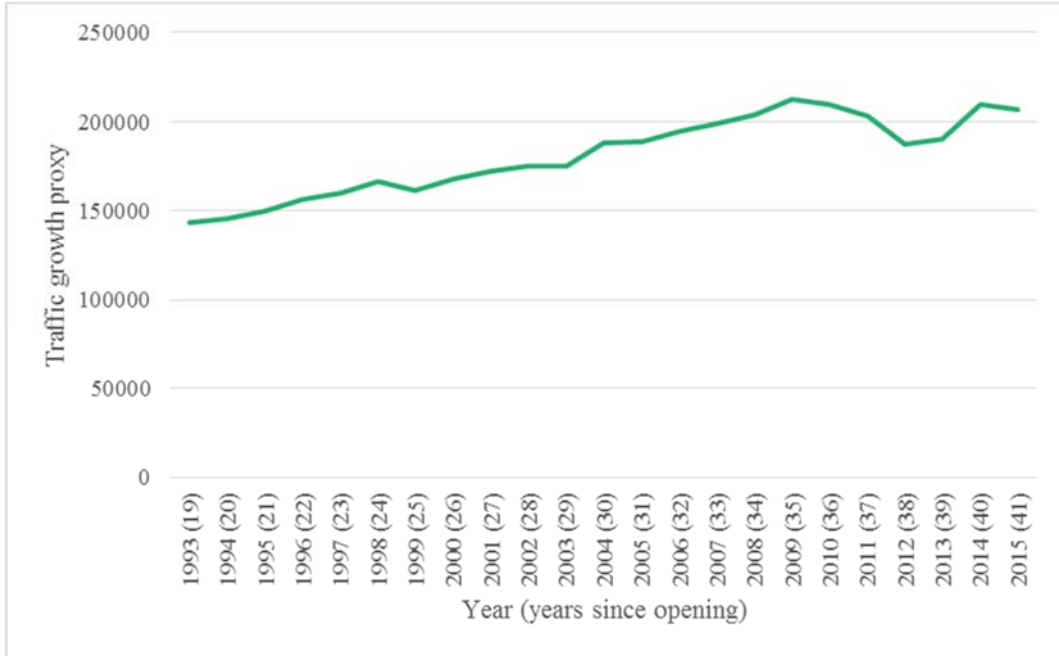


Source: DfT – AADF traffic counts

M5 between Avonmouth Bridge (J18 to J19)

- This section of the M5 was opened in 1974

Figure 87 – Growth in traffic on the M5 around Avonmouth Bridge (J18 to J19)

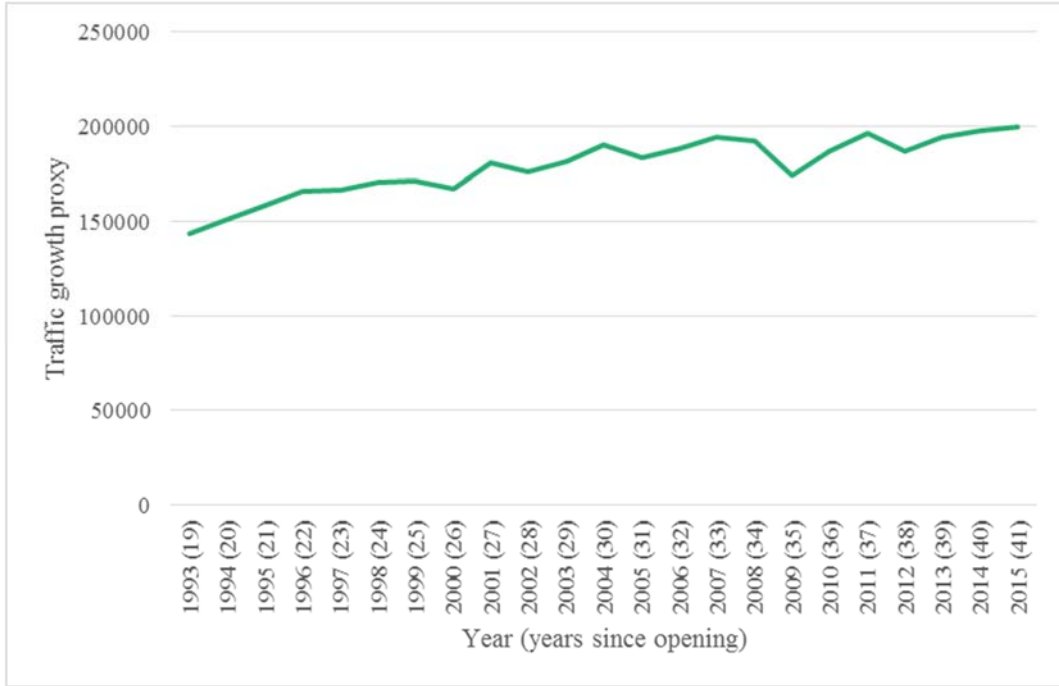


Source: DfT – AADF traffic counts

M40 between Stokenchurch to Waterstock (J5 to J8a)

- This section of the M40 was opened in 1974

Figure 88 – Growth in traffic on the M40 between Stokenchurch to Waterstock (J5 to J8a)



Source: DfT – AADF traffic counts

Appendix B

Road Case Study

B1 Road Case Study

The issues that may arise in the application of both highway and demand modelling when used to determine the estimation of long term demand and benefits in congested networks is considered within this case study.

7.3.1 Highway Model Convergence

Evaluation and appraisal of highway schemes routinely use outputs from assignment models whereby a comparison is made between do-minimum (without scheme) and do-something (with scheme) model scenarios. The comparison is undertaken using the outputs of the assignment models as the product of an iterative traffic assignment process.

In congested networks small incremental flow changes can lead to large exponential changes in delay, resulting in unstable assignment results. If proximity (Wardrop's equilibrium) and stability criteria are not satisfied, then the outputs of the assignment process will not converge, or may lead to differing levels of convergence in the case of the do-minimum and do-something. In both cases, the result being that the true differences between with scheme and without scheme scenarios have not been achieved. This means that comparisons will, across scenarios, result in an overstatement or understatement of true benefits.

7.3.2 Appraisal Example

The congestion and convergence issues described are illustrated in the following example.

7.3.3 Methodology

An appraisal, using the DfT's Transport User Benefit Analysis (TUBA) software, has been undertaken for a junction improvement scheme in northern England. An economic assessment has been based on three modelled years 2018, 2023 and 2033, as is typical following current guidance. In this case study, we include an additional model year to be included within the appraisal. The approach to assessing a horizon of 2060 is described below.

1. Calculate a 2033 to 2060 growth factor. Use this factor to grow the 2033 Do Something and 2033 Do Minimum travel demand matrices to 2060.
2. Assign the travel demand matrices obtained to the 2060 Do Something and 2060 Do Minimum networks respectively.
3. Obtain skimmed generalised cost (time and distance) matrices for both the Do Minimum and Do Something scenarios.
4. Carry out the economic appraisal using TUBA including 2060 as a modelled year (inclusive of 2018, 2023 and 2033 as used previously).

DfT’s “Road Traffic Forecasts 2015” gives traffic forecasts obtained from runs of the National Transport Model (NTM). This document provides forecasts of vehicle kilometres by vehicle type (cars, LGVs and OGVs), in the study area in 5 year increments between 2010 and 2040. In order to extrapolate the growth rate from the last available modelled year (2033) to 2060, the following steps were taken.

5. Use the forecast vehicle kilometres for 2030 and 2040 from DfT’s “Road Traffic Forecasts 2015” to calculate the implied traffic growth rate for each vehicle type.
6. Extrapolate the forecasts to 2060 from the last modelled year of 2033.

Based on the procedure, **Table 18** gives the growth factors applied to the 2033 travel demand matrix to obtain the 2060 matrices.

Table 18 – Growth Factor by Vehicle Class to obtain 2060 Travel Demand Matrices

Car	Light Goods Vehicles (LGV)	Other Goods Vehicles (OGV)
24.43%	51.90%	19.27%

Source: Arup

7.3.4 Impact on Assignment

Table 19 provides example model convergence statistics for one of the modelled periods. This table gives the number of iterations for the algorithm to converge to Wardrop’s equilibrium. As traffic assignment is an iterative process, the next two columns give an indication of the stability of the results; these are the percentage of link flows and delays changing by less than 1%. Finally the last column reports the gap which gives an indication of the proximity of the assignment results to a true Wardrop’s Equilibrium.

This table shows that the inclusion of 2060 significantly increases the number of iterations required to meet a reasonably converged assignment thereby impacting on model run times. It can be seen that when the limit allowed in SATURN (399) is reached in the Do Minimum, the assignment does not reach the same level of stability as in the other years (as measured by the % of link flows/delays changing by less than 1%). Furthermore, it can be seen that the assignments are further away from the true Wardrop Equilibrium as compared to the previous years. Whilst convergence may (eventually) be achieved in the Do Something, the economic appraisal will not be able to distinguish between the effects of the scheme, and the effects of the varying levels of convergence.

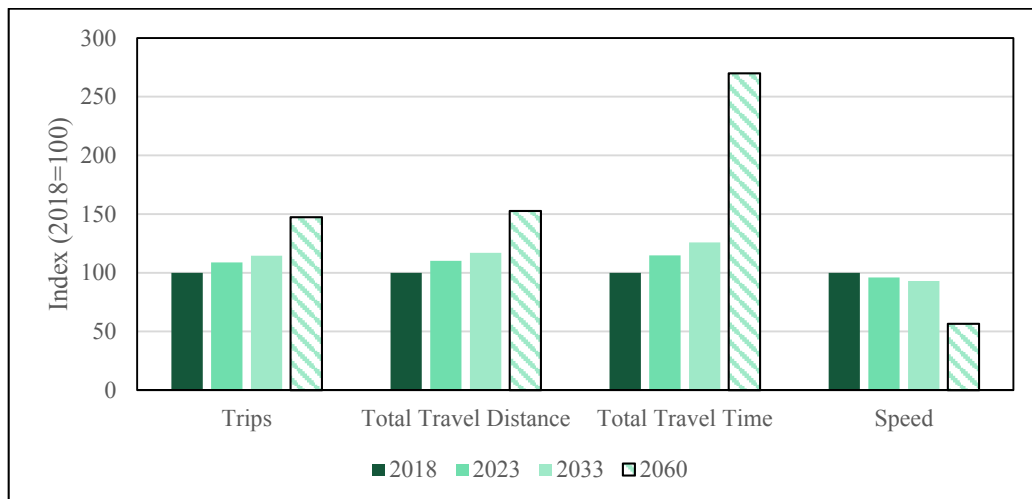
Table 19 – Model convergence statistics

	Number of Iterations	% Link Flows Changing by less than 1%	% Delays Flows Changing by less than 1%	Gap
<i>Do Minimum</i>				
2018	22	98.8	98.8	0.016
2023	25	97.5	98.9	0.017
2033	21	98.3	98.6	0.019
2060	399	95.1	93.9	0.385
<i>Do Something</i>				
2018	23	99.3	98.9	0.007
2023	14	99.5	99.1	0.008
2033	17	98.8	99	0.016
2060	205	99.2	97.9	0.025

Source: Arup

Using the first modelled year (2018) as an index of 100, the number of trips in 2023, 2033 and 2060 are graphed in **Figure 89**. In the same way, the graph also plots the indices of total travel distance, total travel time, and network speeds (relative to 2018). It is evident that the number of trips have increased significantly over this period (trips in 2060 are about 50% higher than in 2018). As users choose routes to minimise generalised travel costs, travel distances and travel time increase. Thus, the network speeds also deteriorate significantly. This points to significant increase in congestion in the network between 2018 and 2060.

Figure 89 – Indices of Trips, Total Travel Distance, Total Travel Time and Average Network Speed by Modelled Year

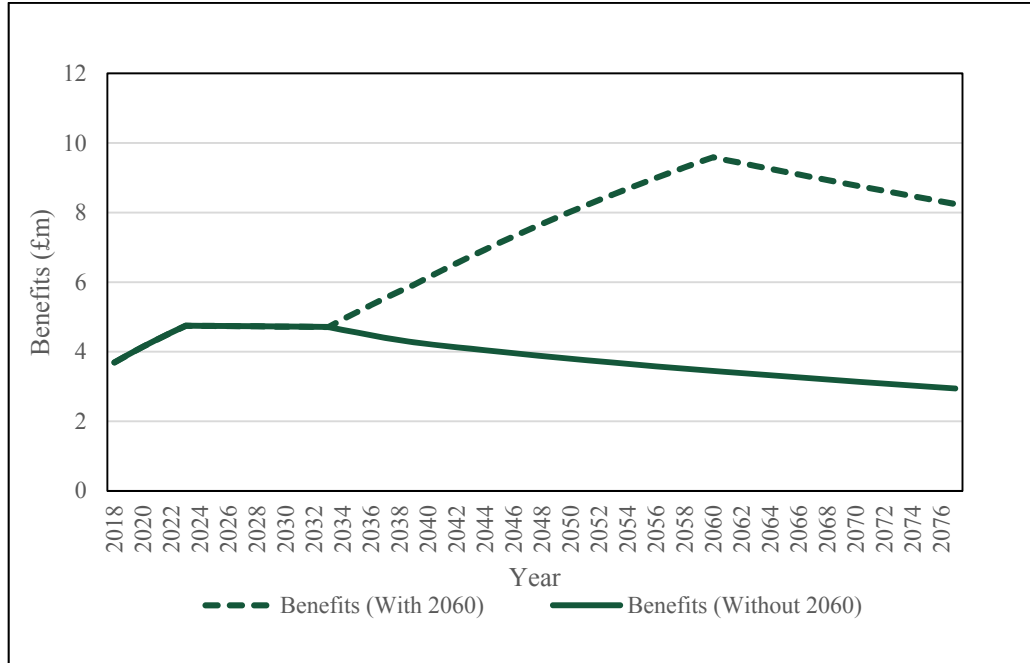


Source: Arup

7.3.5 Impact on Appraisal

The profile of the benefits over the appraisal period, with and without inclusion of the 2060 modelled year is shown in **Figure 90**.

Figure 90 – Profile of Discounted Benefits over Appraisal Period (£m 2010 prices)



Source: Arup

In the ‘without 2060’ case benefits after 2033 decrease as a result of discounting. In the ‘with 2060’ case benefits rise considerably as the delays in the 2060 Do Minimum (that are subsequently removed by the scheme) rise exponentially compared to earlier years. As noted previously, as the model convergence is poor, the substantial 2060 benefits cannot be guaranteed to be completely attributable to the scheme.

On the chart, there are three things happening:

1. congestion grows and, as such, the time savings derived by the scheme grows which drives the benefits upwards;
2. the real value of time increases which drives the benefits upwards; and
3. discounting reduces the present value of benefits by 3.5% per year.

Between 2021 and 2033, (3) counteracts (1) and (2) so the overall effect between 2021 to 2033 is a flat line. However, the benefits rise steeply to 2060 as the size of (1) is big compared to (3). (1) becomes unreliably large in 2060 due to all the reasons stated in the report (convergence, overloading of the network and exponential delay, lack of balance between demand and supply). After 2060, there is no more growth is assumed in (1) or (2) (demand and time savings are ‘capped’ from 2060) and so (3) means that benefits drop.

Table 20 shows the net present value (NPV) and benefit cost ratio (BCR) for this scheme with and without inclusion of the 2060 modelled year. Evaluation with the 2060 model has the effect of significantly increasing both the NPV (close to doubling it) and BCR.

Table 20 – Net Present Value and Benefit Cost Ratio with and without inclusion of Modelled Year 2060

	Without 2060	With 2060
Present Value of Benefits (PVB)	231,541	423,063
Present Value of Costs (PVC)	55,000	55,000
Net Present Value (NPV)	176,541	368,063
Benefit to Cost Ratio (BCR)	4.21	7.69

Source: Arup

7.3.6 Assignment Methodology

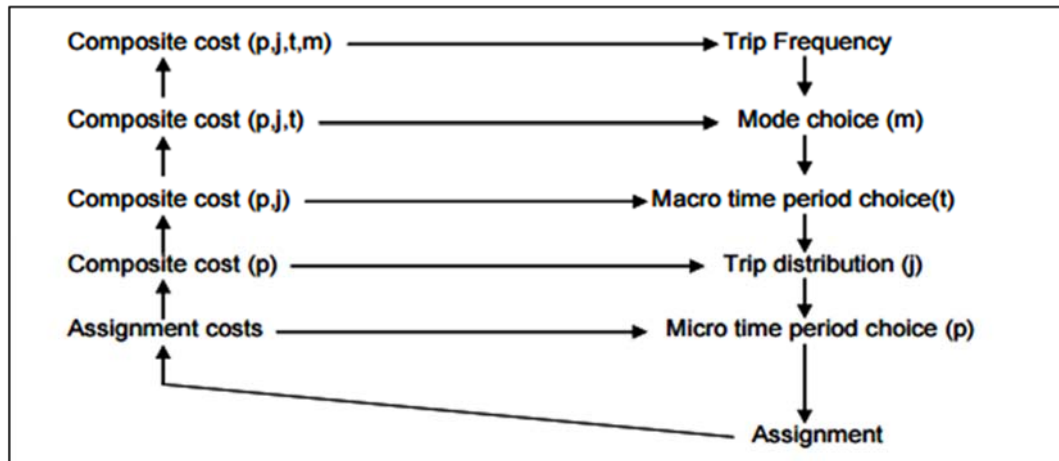
Some assignment models such as SATURN and VISSUM have been extended to take into account for a key feature of congested road networks, which is flow metering.

Flow metering is a procedure for estimating the effects of capacity restrictions on downstream traffic flows such that traffic queued at an upstream junction is not assumed to arrive at a downstream junction within the modelled (usually hourly) period. Failure to take this into account can lead to serious over-estimation of queues and delays at downstream junctions and poor estimation of overall network delays.

Typically Cube Voyager and EMME highway assignment models do not represent flow metering in their standard static highway assignment model software packages. It is anticipated that should the assessment described in the example above have been undertaken using either Cube Voyager or EMME then the overall network delay in the congested 2060 would have been overstated.

7.3.7 Demand Model Issues

The following diagram, **Figure 91**, full VDM process is taken from WebTAG Unit M2 “Variable Demand Modelling”.

Figure 91 – A Typical Choice Hierarchy and Associated Cost Transfers

The least sensitive response (frequency) is at the top while the most sensitive (to cost) is placed at the bottom of the hierarchy. The most sensitive in this context is the choice between routes for an OD pair (i.e. the traffic assignment process).

As described in TAG Unit M2, the cost calculations starts at the bottom of the hierarchy and works its way up the levels, adding one more choice into the cost at each level. The choice calculations are then made down the hierarchy and the whole cycle is recalculated in the next iteration.

As the cost (or more precisely “composite costs”) are calculated as the outcome of an assignment model, poor assignment convergence implies that the costs feeding back up the hierarchy are no longer robust. This means that the composite costs (which are usually log sums when the most common hierarchical logit models are used) are erroneous.

Poorly converged assignments will have implications for the overall variable demand model.

As stated in TAG Unit M2: “A variable demand model includes an assignment stage to provide travel cost information to the demand model. That assignment stage must be adequately converged, particularly since this is necessary to achieve a good level of convergence between the assignment model and the demand model. Assignment can be considered separately from the other mechanisms, but it is essential that an equilibrium solution between demand and supply is obtained.”