COMBINING TIME SERIES RAIL DEMAND AND CROSS SECTIONAL MODE CHOICE MODELS TO FORECAST DEMAND FOR BRITISH HIGH SPEED RAIL

Robert A Cochrane, Visiting Professor in Transport, Imperial College London
Member of the HS2 Analytical Challenge Panel 2009 - 2010

ABSTRACT

The future demand for British mainland long distance InterCity rail services has been forecast for nearly thirty years using mode specific direct demand methods linking demand growth in specific segments (business, leisure, etc) to per capita regional GVA, population, generalised journey time and fares. The elasticities have been estimated using econometric time series analysis of aggregate rail travel data. This is in contrast to urban and sub – regional multi – mode transport demand modelling. The latter has been primarily based on the use of "variable demand modelling" in which the total future disaggregated travel demand by purpose and population segment is first estimated, followed by the estimation of mode choice using nested logit cross – sectional models.

For the lower frequency long distance rail journeys with a minority share of the market, direct demand models have so far proved to be very successful in forecasting aggregate demand and often more successful than multi-mode transport models. They are not suitable for estimating the demand for proposed high speed rail services since they have no means of modelling competition between rail services. A number of "incremental" hybrid models have been developed combining the direct estimation of future total demand by individual mode with the estimation of mode transfer to new high speed services capture through cross sectional nested logit models.

The work has exposed a number of significant problems in combining the two approaches. Two problems appear to affect all the alternative forms of model currently in use. The first is forecasting total long distance travel and in particular, rail travel demand over a period in excess of thirty years. The second is developing mode choice models which can support the financial models for new rail modes such as high speed rail which are needed to allow assessment of alternative fare policies, competition with services on the existing "classic" rail network and the level of subsidy required to operate a high speed rail system.

This paper and the presentation outline current UK long distance rail forecasting and indicate the need for further research in these areas. It is hoped that this will stimulate discussion.

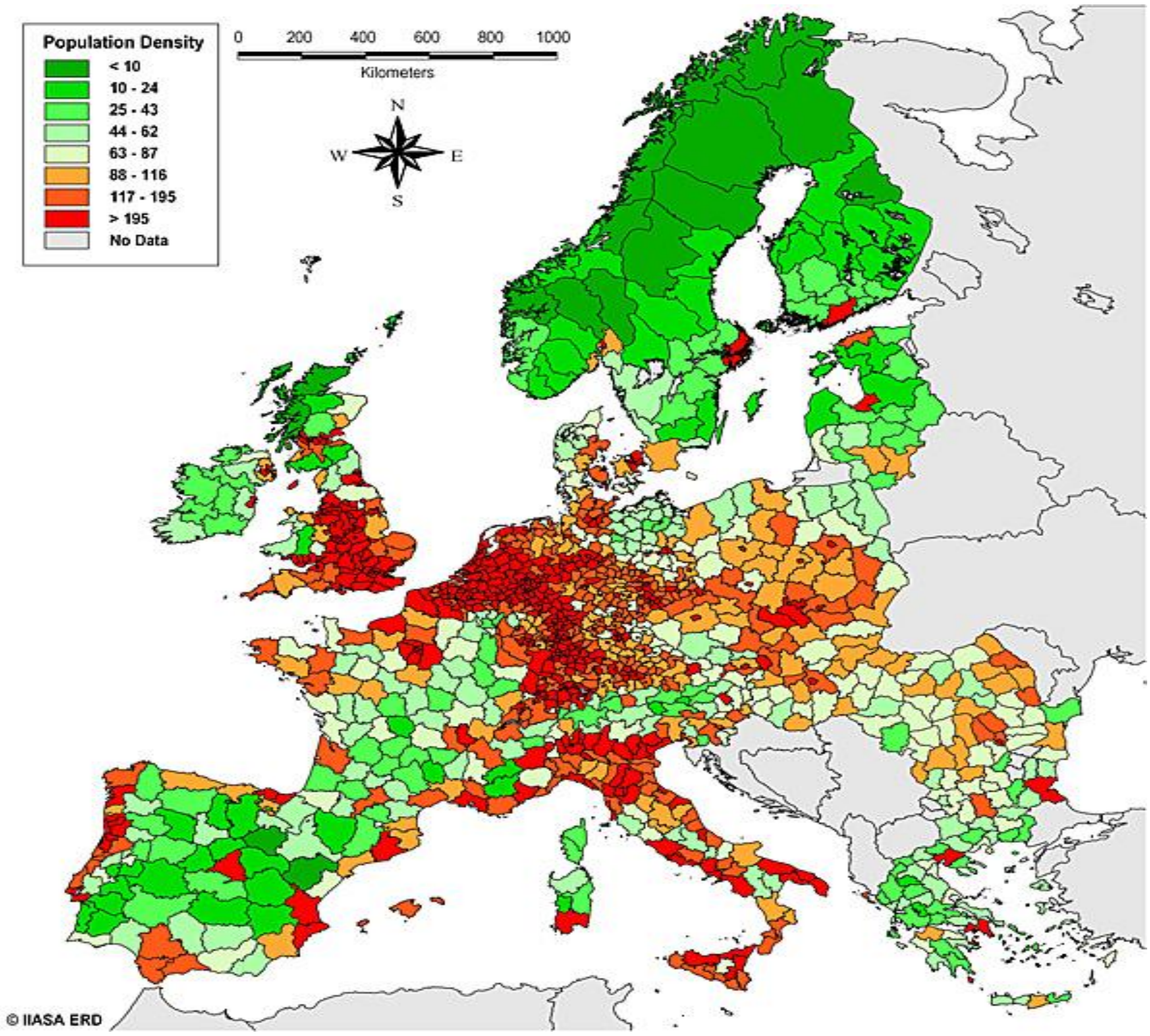
Keywords: rail passenger, long distance, forecasting, demand, mode share, high speed rail

BRITISH LONG DISTANCE TRAVEL DEMAND

Population Distribution and Long Distance Domestic Travel Demand

The largest island within the British Isles archipelago, Great Britain, comprises England, Scotland and Wales. It is roughly triangular, some 1000km from north to south and 600km across the east – west base with a total population of about 60 million. The majority of the population live in a much smaller region, a corridor running from Dover through London and Birmingham to Manchester and Liverpool which is approximately 400km by 200km in extent (Figure 1). As a result, only about 6 million people, or 10% of the population of Great Britain, live more than 400km from the centre of London, the capital. (Figure 2).

Figure I – European Population distribution and densities
(Source: IIASA ERD)

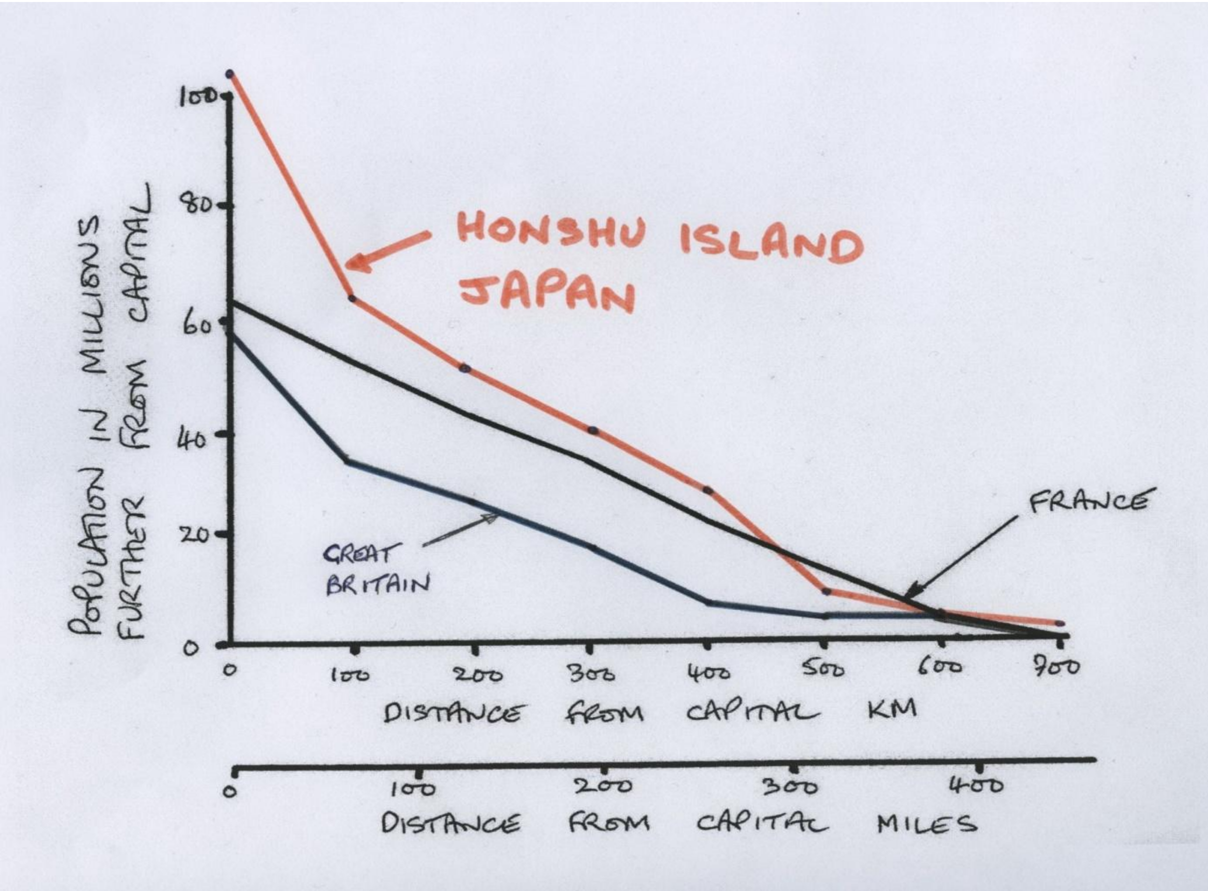


The remaining 10% live mainly in three outlying regions, the Severn Estuary and southern Wales, Newcastle and the Tyne region in the north east and the Glasgow – Edinburgh corridor in southern Scotland.

This concentrated pattern of development is very different from those in other European countries, where a greater proportion of the population is separated by distances in the range of 200km to 600km where high speed rail services are most likely to succeed in attracting passengers from conventional rail and air services.

In common with other developed European nations, British domestic travel is dominated by short distance trips. The major source of data is the ongoing National Travel Survey (NTS). This is a household survey based on diaries kept by about 8,000 households and covering travel by about 19,000 individuals, or about 0.03% of the British population. The survey indicates that the average resident makes approximately three (one way) vehicular or bicycle trips per day and travels about 11,500km each year. The average trip length is about 13km. For brevity, all references to travel and trips in the remainder of this paper are to domestic non – walk travel, including air, within Great Britain unless otherwise noted.

Figure 2 – Distribution of population by distance from the capital; Great Britain, France and Honshu Island, Japan (Source: Cochrane (2008)



In Great Britain, domestic long distance trips are usually considered to be those with lengths exceeding 50 miles or 80km, and all references to long distance trips below will be on this basis, unless otherwise noted. They represent only about 2% of total vehicular passenger trips, but about 30% of total passenger km because of their greater average length which is approximately 180km.

The low proportion of long distance trips means that household random sampling methods such as the NTS provide only limited data sets and detailed analysis and forecasting of total demand is difficult. Nevertheless, the aggregate characteristics of long distance travel and the differences between these trips and the national averages, which are dominated by the shorter trips, are clear.

Table I – Comparison of Journey Purposes for all British domestic trips and Long distance trips (over 80km) (Sources: NTS data tabulated by Scott Wilson and Rand Europe for DfT LDM Studies, Cochrane)

Journey Purpose	All Trips	Long distance Trips
Personal business / shopping	43%	13%
Commute / education	23%	14%
Visiting friends and relatives	15%	24%
Day trip / holiday / other leisure	15%	33%
Business	5%	16%

Table 1 sets out the differences in the distribution of journey purposes as between long distance trips and all domestic trips based on aggregate NTS data for 2005. Two thirds of all trips are journey to work or education together with personal business and shopping. For long distance travel, the pattern is reversed and these trips only amount to 27% of trips. The biggest percentage rises in the proportions of trips on longer distance journeys are business and holidays together with other leisure.

The mode shares also change with journey length. These are set out in Table 2.

Table 2 – Comparison of Journey Primary Mode Shares for all British domestic trips and Long distance trips (over 80km) (Sources: NTS data tabulated by Scott Wilson and Rand Europe for DfT LDM Studies, Cochrane)

Primary Journey Mode	All Trips	Long distance Trips
Car Driver and Passenger	74%	84%
Bus / Coach	11%	5%
Rail	2%	10%
Air	<<1%	1%
Other (metro, light rail, taxi, etc)	13%	1%
Total Public Transport	26%	17%
Average Car Occupancy	1.54	1.68

The percentage of British long distance journeys which individuals make by car rises for long distance trips from about 74% to about 84%. The reason is an increase in the passenger / vehicle ratio from about 1.54 to about 1.68. This reflects in part the use of the car for domestic long distance family holiday and VFR travel.

The relative importance of bus and coach as compared with rail may be surprising to those who live or work in the South East of England, with its high dependence on suburban rail for work, leisure and comparison shopping in central London. Bus and coach are the most important public transport mode based on numbers or trips and provide for the majority of public transport trips and play a particularly important role in providing "safety net" public transport for lower income groups

Perhaps most important, the figures place two frequently expressed opinions into perspective. The first is that improvements in British rail services could provide environmental benefits by reducing dependence on air travel. True domestic air travel is very low. It is confined to origins and destinations poorly served by rail and London to the Glasgow – Edinburgh corridor. The second is the belief that better public transport could replace the use of the car for long trips. Car travel is more dominant on trips over 80km, provided that they are under 560km in length.

These points are examined in greater detail in Table 3, which indicates how the primary mode share for long distance travel changes with distance. In this table, car drivers and passengers have been aggregated for clarity and the average car occupancy is given separately.

Table 3 – Comparison of Journey Primary Mode Shares for Long distance trips by distance band
(Sources: NTS data tabulated by Scott Wilson and Rand Europe for DfT LDM Studies, Cochrane)

Primary Journey Mode	80-120km	120-160km	160-240km	240-400km	400-560km	Over 560km
Car Driver and Passenger	87%	88%	85%	81%	72%	45%
Bus / Coach	3%	4%	5%	6%	10%	5%
Rail	10%	8%	10%	12%	14%	13%
Air	<<1%	<<1%	<1%	1%	4%	37%
Total Public Transport	13%	12%	15%	19%	28%	55%
Average Car Occupancy	1.64	1.73	1.81	1.88	2.00	2.25

Car occupancy rises monotonically with journey length, suggesting a strong link with family or leisure group car use for longer distance trips. The overall car mode share remains above 80% up to 400km, which is similar to the extent of the main British development corridor. This is probably due in part to many trips having origins and destinations within extended urban agglomerations, where the limitation on access to long distance public transport modes is a factor in mode choice, and in part to the nature of leisure travel destinations.

Table 4 analyses the approximate market shares within the long distance public transport sector. The distance segment from 80 to 120km may differ slightly from the general trend because it includes a significant number of long distance rail commuting trips, particularly those in the south east. From 120km to 400km, the range which includes most of the long distance trips within the primary British development corridor, the mode shares are remarkably consistent.

Public transport carries about 15% of all passengers, of which a third (i.e. about 5% of the total) travel by bus or coach and two thirds (i.e. approximately 10% of the total) by rail.

Table 4 – Comparison of Journey Primary Mode Shares for Long distance public transport trips by distance band (Sources: NTS data tabulated by Scott Wilson and Rand Europe for DfT LDM Studies, Cochrane)

Primary Journey Mode – Public Transport Shares	80-120km	120-160km	160-240km	240-400km	400-560km	Over 560km
Bus / Coach	23%	33%	33%	32%	36%	9%
Rail	77%	67%	67%	63%	50%	24%
Air	<<1%	<<1%	<1%	5%	14%	67%

Air travel is not significant until beyond about 400km. At shorter distances, it is the preferred mode for long distance trips where the UK London centred rail network provides poor service standards for business and higher income travellers. Air becomes dominant above 500km. With current long distance rail operating speeds, this represents an in vehicle journey time of about 3.5 hours. This is consistent with experience of rail / air competition in France (MVA 2009, SDG 2009). However, this represents only a very small proportion of British domestic trips, being primarily from the south-east to the Glasgow – Edinburgh corridor.

The coach share remains significant beyond 400km. It tails off beyond 560km, which represents coach journey times exceeding seven hours, or more than a full working day including access time. This seems likely to be a result both of the lower fares charged for coach services as compared with rail, which make long distance travel more accessible to low income travellers. It may also reflect the greater accessibility to coach services for those without the availability of a car to drive to and leave at a long distance rail station or the mobility to cope with access problems in older railway stations.

The volume of long distance travel is highly income dependent. As Table 5 shows, those in the top 20% by income make over four times as many domestic long distance trips as those in the lowest 20% by income. In addition, these higher income trips are more likely to be made by rail than coach for journeys up to 400km and by air or rail rather than coach at longer distances. The switch from rail to air occurs at about 500km, representing an in vehicle journey time by rail of between three and three and a\half hours by fast classic rail.

Table 5 – Comparison of long distance trip generation by income band
(Sources: NTS data tabulated by Scott Wilson and Rand Europe for DfT LDM Studies, Cochrane)

Income Quintile	1	2	3	4	5
Domestic Long Distance trips per person per calendar month	0.8	1.0	1.5	2.0	3.3

Total domestic long distance travel, which is dominated by car travel, has risen by about 15% over the last fifteen years and has been broadly in line with growth in per capita GDP. By contrast, long distance rail travel has grown by about 70% over the same period.

The Independent Travel Commission Report on Long Distance Travel, which was prepared by Professor Dargay of ITS Leeds using NTS data for the period 1995 – 2006 (Dargay 2010), indicates a national total long distance (over 80km) travel growth / GDP elasticity of about 0.35. By contrast, rail travel over the same period (1996 – 2006) shows a long distance rail growth / GDP elasticity of about 1.2 and air travel had an elasticity close to one. The equivalent GDP growth elasticity for coach travel was close to zero.

These figures confirm the conclusions from other analyses, including the National Transport Model (NTM), that long distance travel and in particular, long distance air and rail travel, are highly income dependent and that this dependency has been sustained over the last 15 years without any sign of saturation in long distance travel demand as household incomes rise. (See Dargay 2010 Table 8 for short and long run income elasticities). If carried forward using current long term Treasury forecasts for GDP growth, this results in very high demand for long distance travel facilities over the next thirty years.

Interest in high speed rail demand forecasting has tended to focus on the other leading issue, the forecasting of mode share for a new mode and mode capture from other modes. But forecasting long term total demand for domestic public transport modes appears to be both the most important and the most difficult problem.

Current British Long Distance Rail Technology

Great Britain currently has one high speed rail line (HS1), between London and the Channel Tunnel near Folkestone in Kent. This line also offers medium distance express train services to centres in Kent over distances of 60-100km using "Javelin" hybrid rolling stock operating over the high speed line and the existing conventional tracks. No revealed preference survey data is as yet available for these services and in any event, it would be too early for surveys to indicate the likely long term effect on demand and route choice as the new services, which terminate at St Pancras in London rather than Waterloo, are likely to change origin and destination choices which can lag by up to four years for journeys to work.

The remainder of Great Britain is served by a mixed traffic railway network using a fast conventional technology passenger rail technology sometimes referred to as "classic rail". Inter-City trains use either electric or diesel-electric traction allowing maximum speeds of between 200kph and 225kph. Long distance station to station speeds (including the effect of intermediate stopping points) for express services are in the range of 130kph to 150kph.

This is fast by European conventional rail standards. It is similar to the speed range which the Swiss railways have adopted for modernising their national passenger rail system.

As a result, all major British city centres within the main developed corridor in England, together with the outer developed regions of Cardiff in south Wales and the Newcastle urban and industrial complex are already within about three to three and a half hours of London by rail.

As noted previously, this is reflected in the NTS travel data, which shows rail losing market share rapidly to air beyond about 500km. Detailed studies of travel between British city pairs confirms that the air share of the combined air / rail long distance travel market switches from less than 10% at a rail journey time of two hours to 85% at four hours, the equal mode share point being close to three hours (MVA 2009).

Over the last year, studies have been undertaken by a wholly Government owned company, High Speed 2 Ltd, of a high speed line from London connecting initially with Birmingham and later to be extended to Manchester and through to Glasgow and Edinburgh. The line would cater both for dedicated high speed services and services using hybrid rolling stock which would continue northwards using existing lines, in the same manner as the existing Kent Javelin services operating over HS1.

Alternative Methods of Forecasting Long Distance Travel Demand

The Multi-Stage Transport Model

For the last fifty years, the primary method of analysing and forecasting passenger travel quantitatively in urban and extended urban areas has been the multi-stage transport model, in which travel demand is usually modelled as a four stage process for each travel purpose:

- Estimation of primary demand from exogenous econometric indicators
- Distribution between alternative travel destinations based on relative generalised cost
- Allocation between modes of travel based on relative generalised cost
- Assignment to routes by mode.

As originally developed by Hansen, Voorhees and others in the United States and incorporated in the HUD urban transport modelling package, the approach was heuristic and intended primarily for urban travel demand modelling. Some historic reminders of the origins still exist in terms such as “the Gravity Model”.

Nested mode choice estimation was placed on a sounder statistical foundation within the framework of discrete choice models by McFadden in the early 1970s (McFadden 1972). Perhaps surprisingly, the formal multi – stage logit based discrete choice model was used in the UK for long distance transport modelling at this time and prior to its use in urban travel forecasting. Based on the earlier work of Luce and Quarmby (Luce 1959, Quarmby 1967) and encouraged by discussions with Flowerdew, then at LSE, Cochrane and Pendlebury developed nested mode choice models, estimated using maximum likelihood methods drawn from medical statistics research, for both passenger and freight movements for the UK Channel Tunnel Studies in 1971-1972. (Coopers & Lybrand 1973).

Hence models of this form have been used in the UK for forecasting long distance travel, and high speed rail in particular, for nearly forty years. However, the combination of familiarity and ease of application using fast statistical computer packages which assume an understanding of the underlying assumptions mean that the limitations of these general forms of model, when applied in specific circumstances such as long distance travel forecasting, are not always taken into account.

The first is that they are essentially cross-sectional equilibrium models. Their econometric interpretation is that each traveller behaves as a rational economic decision maker (“homo economicus”), maximising the difference between personal utility and personal cost. They also assume implicitly that perfect information on choices is available to all travellers and that any factors which affect travel have been in place long enough for travellers to have taken them into account.

Human behaviour is more complex than simple econometric analysis assumes. And changes in mode choice may take six months to a year to take place while travellers obtain information, experiment and then settle on new choices. Choices relating to origins and destinations take even longer, up to four years or so, if they are dependent on changes in residence or workplace.

Because they are cross-sectional models, forecasting future travel requires time series forecasts of a very wide range of exogenous variables, not only to forecast demand, but also to provide appropriate times and costs for the supply side forecasts. In theory at least, the longitudinal and cross- sectional models can be combined in a form of a panel data model, but the computational problems of model estimation and the data requirements preclude this approach except under special circumstances or with spatially or temporally localised models. (For a review of panel data models, see Greene 2008 Chapter 9).

Forecasts are normally made using "pivot" or incremental techniques, in which the synthetic model results for the base year are adjusted to equal the base year data. The most commonly applied technique is the ratio method, although the difference method may be used where extreme ratios are encountered, as with empty cells in the base year data.

The majority of discrete mode choice models assume that mode choice may be modelled using the Gumbel extreme value distribution for generalised cost, with a linear generalised cost vector.

It is not always appreciated that with the exception of the distribution stage, where Cochrane showed that the Gumbel extreme value distribution arises naturally from choice between a large number of alternatives (Cochrane 1976) this is an arbitrary choice made because of its convenient statistical and economic properties and in many circumstances only provides an approximation to actual travel behaviour.

Aside from the more complex and often intractable probit model form, one alternative is to use the Weibull distribution, which is the logarithmic analogue of the Gumbel and has the same relationship to the Gumbel as the log-normal does to the normal. This is also equivalent to using a log transformation of the generalised cost equation.

In the logit form, the standard discrete mode choice model, mode shares depend on generalised cost differences, so deducting a uniform cost from all modes does not alter mode shares. Under a logarithmic transformation, mode share depends on generalised cost ratios. The first form has the advantage that common generalised cost elements do not affect choice probabilities, the second that the model parameters (which are probabilities and elasticities) are unaffected by changes in the scale of utility measurements and hence money measures are unaffected by uniform inflation. Related differences exist when comparing elasticities.

The most common practice in current long distance modelling is to use the logit model for the error function, but to test logarithmic forms for individual generalised cost components. This may improve the model fit for one of several reasons:

- There is evidence that the greater length and range of distance and time lead to non – linear relationships between out of pocket cost, time and the perceived generalised cost. (Gunn and Pedersen 2007)
- The forecasting of demand for modes with low market shares is likely to be affected by the tails of the distributions adopted for modelling dispersion
- Longitudinal time series models, with which these models are compared when assessing rail forecasts, almost without exception use constant elasticity of substitution forms.

A third issue is that these multi-stage models require large volumes of data across a range of modes and are dependent on household survey data which is difficult and expensive to collect. This is particularly the case if a wide range of modes with differing levels of penetration and a large range of exogenous variables are to be taken into account. This is a serious problem in long distance travel analysis and forecasting, where data is very limited.

As a result, these models may not be as successful when estimating incremental future mode shares for a minor mode, such as rail in Great Britain other than for passenger travel to Central London. Obtaining sufficient data would also be a very expensive means of separately developing forecasts for a minor mode in the absence of a joint multi – modal travel forecasting programme.

The final problem is one which is common to all types of model when forecasting modes such as high speed rail; the appropriate treatment of "new" modes. Issues include establishing the generalised costs of a new mode and the appropriate measurement of user benefits (Cochrane 2009a).

The generalised cost characteristics of the new mode are first evaluated in terms of the characteristics used for existing modes such as time and cost. This allows forecasting to be based as far as possible on factors such as the values of time and cost which have previously been estimated for existing modes using revealed preference survey data.

One practical difficulty is forecasting the total origin – destination times and costs for the new mode, particularly when, as is usually the case, a new and faster mode makes fewer stops, thereby becoming more dependent on access time and cost. Paradoxically, accurate forecasting of demand for long distance high speed rail requires accurate modelling of the local road access networks.

A second difficulty is that proponents of a new mode may claim that it is so fundamentally "different" in its improved speed, reliability and comfort that it can only be modelled using a "mode constant" which should be added to other factors in the generalised cost. The effect on mode share is dependent on whether this mode advantage is considered to be a perceived premium per trip or whether it is dependent on journey length or in vehicle time.

A review of past studies of the introduction of progressively faster and more reliable diesel and electric long distance passenger services suggests that any "mode constant" found in early studies can usually be replaced in the long run by more fundamental generalised cost factors common to other modes.

Another issue in forecasting demand for a new mode within the multi-stage model is to determine its appropriate position within a multi – mode choice hierarchy. For long distance modelling, this problem goes back to the 1972 Channel Tunnel Studies, in which the French regarded the Tunnel as a completely new, alternative travel choice for travellers between Great Britain and Continental Europe, competing directly with all sea services, whereas the British analysts considered it to be a new alternative "route" competing with individual combinations of rail and sea ferry. The alternative assumptions gave differing forecast mode shares.

This form of structural difference persists in current long distance models, which differ in the nesting of long distance travel modes. For example, The Planet Strategic Model (PSM) as modified for HS2 forecasting uses a classic rail – high speed rail nest at the lowest levelof its mode structure, with air competition modelled at a higher level. The DfT Long Distance Model (LDM) also includes air in the lowest level public transport nest.

This difference would not have a significant effect on trips under about 300km, where the air share of travel is negligible, but it is not clear what effect it might have on forecasts for the London – Scotland market, where air is and will continue to be a serious competitor for domestic traffic.

As a result of the UK land use pattern, modelling the direct competition between HSR and classic rail is the most critical factor in setting an appropriate fare premium for the higher standard of service which will minimise the subsidy required for infrastructure investment if not also for operations. In the absence of revealed preference data derived from comparable situations elsewhere, the only possibility is up to date stated preference interview surveys.

Stated preference surveys raise specific problems, some of which have been encountered in British high speed rail surveys and which are discussed in more detail below. This is the second critical area in HSR forecasting, alongside total demand forecasting, to which we will return below.

Direct Demand "Incremental elasticity" Models

The alternative form of model to the standard multi – stage transport model for forecasting rail demand is the direct demand, or incremental elasticity model. This form of model estimates growth in demand between individual origins and destinations, or groups of origins and destinations, for a single mode, using elasticities obtained from time series analysis data primarily drawn from mode specific sources such as ticket sales.

These models are founded on a base year matrix of travel which can then be forecast forward using elasticities which relate to changes in per capita income or a suitable surrogate, fares, rail travel time, and comfort factors such as crowding. In combination with a network model, changes in routing resulting from changes in the times and fares on individual routes can be estimated provided they have the same station origins and destinations. Since the ticket information does not include data on the actual origins and destinations of trips with a rail leg, switching between stations (for example, onto a faster train with limited numbers of stops) cannot be modelled.

Neither are the journey purposes, which have different fare and time elasticities, known directly from ticket information. These must be estimated by ticket type from separate surveys, taking account of the very significant differences in travel purpose by journey length exposed by NTS data.

An important limitation is that the fares elasticities do not take account of internal rail competition. They are long run and assume that where fares change on a rail network, all fares change in parallel, thereby cancelling out any effects of competition between rail fare classes or routes.

The great attraction of these models is that they have over the last fifteen years provided more reliable medium term growth forecasts than multi-stage models. The major reason is that mentioned earlier; whilst total travel measured in passenger km has increased by about 15%, rail travel has increased by about 70%. Back – casting exercises using currently available multi-mode cross – sectional models such as the National Transport Model (NTM) have found difficulty in explaining this growth synthetically, whereas simple incremental demand forecasting models based on ticket sales and elasticities developed from time series for individual modes have been very successful.

The limitation on using these direct demand models is that in most cases, simple longitudinal time series cannot distinguish easily between exogenous demand increases resulting from effects such as income changes and inter-mode competition due to changes in (for example) long term changes in the level of car ownership or the competitiveness of low cost airlines. PDFH includes advice on cross – elasticities between modes, but these estimates are highly aggregated and less reliable than the direct elasticities. As Dargay points out in the ITC Report, cross- elasticities are not only less reliable than direct elasticities but are also in many cases quite low. Changes in the relative cost or time of competing modes produce relatively low switches between modes. (Dargay 2010).

The greatest weakness of these models for forecasting the demand for high speed rail is that they have no means of modelling either direct price or direct service competition between railway routes with significantly different routes, locations and frequency of stations, speeds and fare structures. Neither can they forecast the effect of rail improvements on the road system, either those resulting from reductions in overall car travel or greater congestion close to rail terminals. This also means that they cannot provide direct estimates of changes in the generation of greenhouse gases due to rail system improvements.

They remain important because of their role in forecasting exogenous rail demand when using hybrid models, such as the PLANET strategic model used for HS2 high speed rail which incorporates separate long term forecasts of rail and road demand. They also provide valuable benchmarking results for checking the performance of cross- sectional models in forecasting demand for minor modes.

Current Methods of Forecasting British Long Distance Rail Demand

Since 1986, rail travel demand in Great Britain for suburban, rural and long distance services has been forecast using the direct demand methods set out in the Passenger Demand Forecasting Handbook (PDFH). These were originally developed by econometric research groups with the British Railways Board headquarters and supported by consultants prior to rail privatisation. Ongoing development is now carried out under the Passenger Demand Forecasting Council (PDFC) and managed by ATOC, the association of train service operators. The Fifth Edition (PDFH 5.0) was released in August 2009.

The Handbook is a private and commercially confidential document, only available on annual subscription to members and associate members of PDFC. The Department of Transport is a member of the PDFC, but it takes an independent line on many matters. This is understandable and appropriate, since approximately half of the cost of operating the national railway services comes from a public subsidy which in all forms is currently of the order of £5 billion per year.

The DfT advice is set out as amendments to the PDFH in the form of Transport Advisory Group (TAG) guidance notes which are available on the web as WebTAG Unit 3.15.4. These guidance notes are in the public domain, but they are not always easy to understand in full without access to the original PDFH reports.

The draft TAG guidance issued in January 2010 following release of PDFH 5.0 has rejected a number of important changes from the previous PDFH 4.1 pending further research and requires the earlier edition to be used in the meantime. In the case of long distance travel, the most important differences relate to:

- Journey purpose / ticket type allocation
- Long distance travel growth elasticities
- Fares elasticities

Rail Journey Purpose / Ticket type allocation

Origin / destination information is obtained primarily from the LENNON rail ticket sales database which is used as a proxy for rail demand and is considered to be commercially confidential. PDFH 5.0 uses a new and explicit mapping of ticket types to journey purposes, derived from National Rail Travel Survey (NRTS) data. Unfortunately, the classifications adopted are not consistent with those used in collecting NTS data and for good reason are not supported by the DfT. DfT has now issued its own NTS – compliant correspondence tables. These latter are essential for combining rail ticket and NTS data in hybrid models incorporating both rail direct demand modelling and cross – sectional road models.

Long Distance Rail Travel Growth Elasticities

Long distance rail demand is currently forecast indirectly from forecasts of regional gross value added per capita (GVA/capita). Regional GVA is regional GDP less taxation on production plus subsidies and GVA/capita is essentially a surrogate for regional average income change.

Table 6 – Comparison of GVA /capita long distance rail trip generation elasticities in km
(Source: Draft TAG Unit 3.15.4 guidance, Cochrane)

Journey Direction	PDFH 4.1 GVA/capita elasticity (Current DfT Advice)	PDFH 5.0 GVA/capita elasticity (Current PDFC Advice)
To London	2.00 + 0.002 per km	1.9
From London	0.84 + 0.002 per km	0.9

The rail growth forecasts are obtained by applying the elasticities which have been obtained from time series analyses carried out by ITS Leeds. It is understood that both sets of elasticities\ were estimated from the same data set, but that the variation with distance was not included in the PDFH analysis. The figures quoted in WebTAG are shown in Table 6 above, converted from miles to km.

It is widely recognised that for distances of over 200 miles (325km), the PDFH 4.1 elasticities are unrealistically high. For trips from London to Edinburgh, the PDFH 4.1 elasticity is about +3.3 for trips to London and 2.2 for trips from London. These may be compared with that in the ITC Report (Dargay 2010) of about 1.2 for total long distance rail travel.

Interestingly, the MVA Report (MVA 2007) which formed the basis for the development of the PDFH 5.0 recommendations suggested rather different figures based on unreferenced work by ITS Leeds. These indicate a fall in the elasticity with distance for journeys to London from 2.0 at 170km, down to 0.7 at 460km and -0.6 over 570km. For journeys from London, the corresponding figures range from 0.7 at 170km, down to 0.2 at 460km and -0.4 over 570km.

A review of possible data sources suggests that elasticities derived from ticket data may well be strongly affected by an increase in air competition from low cost airlines over the last fifteen years, particularly for journeys between the South East and Scotland. The NTS data, which could potentially separate the positive growth and negative competition elasticities, may be too sparse to quantify any distance effect with precision except by aggregating data over longer time periods.

A further limitation is that the PDFH and WebTAG guidance long distance growth elasticities do not distinguish between income groups or travel purposes. Dargay (Dargay 2010) confirms the importance of these two factors in forecasting long distance travel growth, both in total and by mode.
The conclusion must be that the use of single mode rail direct demand models with elasticities based either on the current PDFH 5.0 advice or the current (interim) WebTAG guidance based on PDFH 4.1 is not adequate for long term forecasting of long distance rail demand, even in the absence of a major intervention such as high speed rail.

Rail fares elasticities

The PDFH fares elasticity chapter has been restructured using base elasticities which are intended to be interpreted as long run elasticities, conditional on all rail fares moving in line (i.e. with a uniform ratio) and that other mode fares remain unchanged. As in the case of demand growth, the DfT recommends the use of the PDFH 4.1 structure and elasticities strictly on an interim basis, principally because it is not clear to them that the methods used in the primary research support this interpretation. PDFH 5.0 contains no separate elasticities for very long distance trips and recommends that full multi modal models are constructed to model competition with air on time and (presumably) fare.

The base figures for long distance rail trips to and from London (strictly, from outside the South East to the London Travel card area, which would correspond approximately with a trip longer than about 50km) range from -0.75 for season tickets and commuting, through -0.8 for business to -1.25 for leisure.

Unlike the demand growth elasticities, which drive overall rail growth in the current high speed rail models, these elasticities are merely indicative benchmarks for HSR forecasting since the major components of HSR demand are capture from classic rail and (depending on the fares policy selected) additional generated travel. The direct HSR fares elasticities and the cross elasticities as between HSR, classic rail and air can only be properly assessed using a cross – sectional model.

Rail travel time elasticities

For many years PDFH has used a generalised journey time (GJT) to assess the effect of increased speed on rail demand. This is essentially the in vehicle station to station journey time, to which service interval and interchange penalties are then added. Demand changes are estimated using elasticities applied to changes in the GJT. PDFH gives a uniform elasticity in demand for all trips between points outside the South East and the London Travelcard region of -0.9.

The Handbook states categorically that there is no evidence to support different GJT elasticities by journey length, direction of travel or travel type. There is no comment on these figures in current WebTAG guidance, but several studies have found elasticities for small changes in journey times to be of this order.

Again, these figures can only be considered as benchmark comparisons for studies of HSR introduction. They may be slightly too low for this purpose. Experience in France suggests that earlier TGV lines had in vehicle time (IVT) demand elasticities up to about -1.6, although more recent lines show figures of the order of -1.0. There is evidence from studies of the introduction of electrification and high speed diesel services in Great Britain over a period of forty years that the IVT elasticity for major journey time improvements is higher and is in the region of -1.1. Note that on theoretical grounds, slightly higher demand elasticities are to be expected for IVT, the difference depending in part on the importance of frequency.

High Speed Rail Forecasting Models

Two types of model are currently available or under development which have the potential to provide forecasts of future high speed rail demand. These are multi – mode models and "hybrid" models which forecast aggregate transport demand growth separately for each mode (air, rail, road private transport) before estimating changes in mode which might occur as a result of changes in cost, speed or other variables on each mode.

Models used for long distance travel forecasting in Britain include the DfT National Transport Model (NTM), the PLANET Strategic model and (currently in the development stage) the DfT Long Distance Model (LDM).

The National Transport Model

The DfT National Transport Model is a very detailed national model covering both short and long distance travel. Total transport demand is forecast in considerable detail by purpose, household structure income and car ownership. A destination and mode choice model estimated using NTS data is then used to forecast total travel demand by origin and destination and by mode.

The modes covered include walk, cycle, car, driver, car passenger bus rail, metro and taxi, but not at present air or high speed rail. Mode choice is based on distance band and zone type, using values estimated from NTS data which are compared with WebTAG recommendations for values of time etc by mode and distance band. The supply side is represented by detailed road and rail networks, and the model is run iteratively for each forecast year and policy option until equilibrium is achieved.

The model could in principle be used to forecast high speed rail demand by incorporating additional routes with appropriate service speeds and connections into the rail network. Its ability to analyse multiple options is limited because it was developed to test national transport policy alternatives rather than specific alternative network infrastructure interventions and it is not designed to allow quick and easy coding of alternative networks.

The effects of competition between high speed and classic rail could be modelled by fare and time differentials between the alternative routes leading to changes in route choice, but the model does not have a probabilistic mode choice stage which could model choice between high speed and classic rail modes in detail. Neither is it linked to the separate UK air forecasting model, the reason in this case being that the model is primarily intended to assess English surface mode transport policy, regional transport policy in Scotland now being a devolved responsibility.

In addition, the level of detail needed to model trips at all scales from walk to long distance, means that model run times are relatively long and it would not be easy to analyse the range of options needed to establish and optimal HSR route structure within a tight project timescale.

The PLANET Strategic Model used for High Speed Line 2 (HS2) Forecasts

The HS2 studies carried out by Atkins in 2009 – 2010 used a development of the PLANET Strategic Model. This model was originally developed in 2002 by Atkins and Arup for the Strategic Rail Authority (SRA) as a means of evaluating potential improvements and additions to the national rail network. (Atkins / HS2 Ltd 2010). As used in the HS2 studies, it incorporated PLANET Long Distance as the primary long distance multi – mode model, regional PLANET models of rail services in the Midlands and the South East and a spreadsheet airport access model. The regional PLANET models rely on elasticity techniques similar to those used in PDFH for local rail modelling.

The choice of model for the HS2 studies was partly dictated by the extremely tight study time constraints, which made it virtually impossible to carry out additional survey work, develop a new model form or extend an existing model more than the minimum required to fulfil the remit.

In common with the majority of current British HSR models, long term travel demand is separately forecast by mode using DfT forecasts by mode for road, rail and air. Forecasts are prepared from these sources for three purposes: business, commuting and leisure.

The unrealistically high GVA related demand elasticity for very long distance rail trips forecast on the basis of PDFH 4.1 is capped to a figure of 2.8, still far above the figures suggested by recent ITS studies. Total demand is capped again at 2033, which, on the basis of the selected elasticity, is close to the date when the initial infrastructure is likely to reach capacity.

These individual forecasts are then combined in a PLANET Long Distance multi – stage mode choice model to forecast the effect of changes in generalised costs by mode.

The stages in the HS2 model (taken from the top) are:

- All Public transport – private transport (i.e. car driver & passenger)
- All rail – air
- Classic rail – high speed rail

Coach is not included as it was not considered that coach travellers would be likely to switch to high speed rail. This is a realistic simplification, given that the remit was not to examine national long distance travel policy for all income groups but a specific proposal to build a high speed railway.

The model does not include a distribution stage to forecast changes in destination choice as relative travel costs change. Forecast trip matrices are prepared by factoring base year matrices using heuristic scaling methods to ensure that the marginal origin and destination figures are equal to the forecast totals.

It was not possible to extend the existing model within the study period, but this is an important potential weakness which should be considered in further studies. Since high speed rail has fewer stations and selectively increases long distance accessibility to their neighbourhoods, re-distribution should expected to occur.

In 2003, the classic rail – high speed rail nested stage form was based on a stated preference survey carried out in 2002 for the SRA and its subsequent analysis. This has been criticised for its analytical structure, which sought to include a poorly defined "mode constant" differential between classic rail and high speed rail. It is also possible to argue that the results may have been biased because it was carried out during the period of slow main line service running following the fatal accident at Hatfield, both of which could have affected attitudes to classic rail. In addition, the survey is now out of date, both because of its date and because many passengers will now have experience of both current British classic rail service standards and high speed rail in England as well as in France and Belgium.

Based on the currently available technical reports, the final model used for the HS2 studies appears to have dropped the earlier mode constant common valuations within the generalised cost for time a fare cost as between the two rail modes. A form of "mode constant" was included only to allow for increased reliability on HSR, estimated as a reduction in average late arrival and valued at three times in vehicle time, as is assumed for waiting time at interchanges.

The model appears to be a pivot (incremental) model, but it is not clear whether the ratio method or the additive method was used to adjust synthetic base year results to the base year figures.

Similarly, it is not clear how the "composite" all mode generalised costs were used to forecast generated travel. This is a major element of the forecast HSR demand since the central forecasts are based on the assumption that, no premium fare would be charged for high speed rail over and above existing classic rail fares. Comparison with WebTAG / PDFH 4.1 parameters suggests that in the absence of any change in fares, the PDFH 4.1 elasticities for generated traffic with respect to time savings were used. These are similar to, and perhaps slightly below historic British Rail figures for major service improvements (see above).

A weakness of the PLANET Strategic rail modelling framework which is recognised in the technical report is the simplified road model and network. This can affect HSR forecasts in two ways. First, as noted earlier, a good road network is needed to ensure that access to the limited number of high speed railway stations, some of which will be in central city areas, is correctly modelled. A good roads network is also needed to assess the impact of rail interventions on the road system, particularly as these will be small in road traffic percentage terms.

The most serious ongoing weaknesses of the HS2 model appear to be a result of the time scale of the study, the narrow remit and the choice of fares policy for the central forecasts. This is the absence of both any detailed segmentation of travellers by income group and a detailed generalised cost model within the lowest rail – rail nest which would allow the impact of differential fares ("premium" fares) as between high speed rail and classic rail on relative mode shares.

In purely technical terms, this was not required by the remit. But it is clear that further studies will of necessity have to consider issues of competition with private sector classic rail operators, potential private sector finance for maintenance, rolling stock and operations and inter – group, the affordable level of subsidy during a period of national austerity and equity. Whilst it is undoubtedly true that lower income groups do use high speed rail, it is a fact that higher incomes groups travel long distances more frequently and are more likely to use modes such as air and high speed rail.

These issues lie outside the scope of the present paper. The point to be made is that the current PLANET model does not have any mechanism which would allow these important issues to be addressed. Moreover, the age, limited sample size and potential for bias in the earlier SRA survey mean that an updated stated preference survey is essential to construct a suitable nested sub – model for the rail modes.

The DfT Long Distance Model

A DfT team, supported by Scott Wilson and Rand Europe, is currently developing a long distance travel model which has not as yet been reported other than at “work in progress” presentations (Rohr et al, 2009) but which could potentially overcome most, if not all, these weaknesses.

The model is a multi – mode long distance domestic travel model covering all Great Britain trips over 80km (50 miles). Total travel is first forecast using time series analyses. It is then allocated to modes using a cross sectional multi level model in which both rail modes and air are included in a single public transport nest. Both road and rail networks have been developed, the road network being in sufficient detail to model station access.

The models of travel demand generation and choice between existing modes have been developed and estimated using NTS data, aggregated over several years where necessary. A new stated preference survey has been carried out to provide data on which to base the choice as between existing classic rail and high speed rail.

There have as yet been no reports as to the performance of the model in the two most difficult areas – the ability of the mode choice model to forecast the effects of both in vehicle time differentials and fare differentials between rail services and the performance of the overall long term travel demand module.

If the stated preference survey has a large enough sample and has been well structured, there seems to be no reason why it should not provide a major improvement in modelling the mode share as between classic and high speed rail. This ability is dependent, as noted above, on the road network and road speeds being adequate to model rail and airport access as well as long distance car travel. Since the majority of the road traffic is making trips of less than 80km, modelling road speeds effectively is an important factor.

A model of this form can form the foundation for a proper competition analysis, financial analysis and subsidy appraisal.

The ability of the model to forecast the growth of minor modes over time will depend on both the time series forecasting of long distance travel and the performance of the mode choice model. The conceptual framework has been in place for forty years. The costs of individual modes can be combined and folded back to model higher levels of mode choice, particularly public versus private travel, and a distribution stage. The highest level logsum then represents an accessibility measure (Cochrane 1975 & 1976, Koenig 1975) which can be used as an additional demand generation factor. In forecasting this process is then be reversed.

The challenge for the LDM team is to show by retro-forecasting that that the large differences between long distance rail travel growth, domestic air travel growth and total long distance travel growth over the last fifteen years can be explained using the standard methodology.

Total Travel Demand Saturation

If these primary goals, which are essentially those of establishing a\ conventional multi-stage forecasting model using the limited data currently available, can be achieved, we will be left with the final forecasting mountain to climb – the question of possible saturation in the long term demand for long distance travel.
This is important in itself, because primary demand growth is the fundamental driver of long term investment in transport infrastructure, but also because of the nature of UK national cost benefit analysis. This uses relatively low discount rates (3.5% reducing to 3%) and very long periods of evaluation. Whilst this may be justified in part on the basis of inter-generational arguments, in combination with a low hurdle figure for the benefit – cost ratio (strictly, the benefit - subsidy ratio for public projects) it creates a portfolio of projects far in excess of available public finance.

One reason is that GDP related demand growth elasticities of over one, coupled with GDP growth forecasts of over 2% lead to very high long term benefit streams. The question is whether the elasticities of demand growth with GDP measured over the last fifteen years are sustainable over the next twenty to thirty years.

The question is closely related to the allocation of household expenditure. Studies of household expenditure suggest that the proportion of the household budget allocated to transport (which is, with minor exceptions such as very long distance commuters, predominantly accounted for by expenditure on the car) has remained relatively constant over the last five years. Studies of the income elasticity of travel expenditure also suggest a figure close to one. Dargay's report for ITC (Dargay 2010) is currently the most helpful source. She notes that long distance travel increased more rapidly than short distance travel from 1996 to 2002, but less rapidly thereafter. Later NTS data from 2007 and 2008 suggests that per capita distance travelled by all modes (and distances) declined from 2005 onwards. Dargay is reluctant to see any trend in these figures, but it is possible that we are coming to the end of a period of twenty years expansion in the growth of long distance domestic travel.

Britain is a relatively small offshore island and domestic long distance travel competes within a wider travel market including international travel over much longer distances for both business and leisure. So within a wider discretionary household travel budget, it may be possible that the overall domestic travel budget will continue to expand with wealth, but the long distance domestic budget may indeed be starting to saturate. More research in this area, and more data is needed.

Conclusions

There are a very wide range of factors which must be taken into account in forecasting long distance travel and high speed domestic rail travel in particular. The most important are those at bottom and the top of the multi – stage travel forecasting pyramid.

At the base, we need robust models of inter – mode choice which can model fare differentials and support financial and competition models. At the top, we need robust models of total future travel growth which can relate travel demand to household income and GDP growth and which can include the effect of possible saturation in the demand for a specific good – long distance domestic travel.

And in conclusion, it is interesting to note that the primary importance of these two factors is unchanged since these forecasting techniques were first used for long distance travel analysis nearly forty years ago in 1970 – 1973 on the Channel Tunnel high speed rail project.

References

McFadden (1972) "Conditional Logit analysis of Qualitative Choice Behaviour"
Luce (1959) "Individual Choice Behaviour" Wiley
Quarmby (1967) "Choice of Travel Mode for the journey to work – some findings" JTEP
Harris & Tanner (1974) "Transport Demand Models based on personal Characteristics", TRRL Supplementary Report 65UC
Gunn & Petersen (2007) "Alternative Strategies to estimating large scale travel demand models" European Transport Conference, Noordwijkerhout
Neuberger (1971) "User Benefit in the Evaluation of Transport and Land Use Plans" JTEP
Atkins (2003) "High Speed Line Study Milestone 6 Report"
Atkins(2003) "High Speed Line Study Milestone 7 Report"
Dargay (2010) "ITC Report on the Prospects for Long Distance Travel in Britain" Joyce Dargay, ITS, Jan 2010, available from the ITC website)

Cochrane (2008) “European Long Distance Travel – Rail system selection and Population Distribution” Submitted to the HS2 Challenge Panel, June 2009
Cochrane (2009a) “Models of Mode Choice and Implications for HS2 Forecasting and Appraisal” Submitted to the HS2 Challenge Panel August 2009
Cochrane (2009b) “Growth in demand for UK Long distance Rail Travel” Submitted to the HS2 Challenge Panel, September 2009
DfT (2010) Draft Rail Demand Forecasting Methodology, TAG Unit 3.15.4” January 2009 (Available on the web through WebTAG)
PDFC (2009) Passenger Demand Forecasting Handbook Version 5.0 (Private document, not available except by subscription)
HS2 Ltd and Atkins (2010) “HS2 Demand Model Analysis” February 2010 (Available on the DfT website)
Mackie Owens & Phillips (1987) “The characteristics of railway passenger demand” JTEP
Coopers & Lybrand (1973) “Channel Tunnel Economic Report” British Channel Tunnel Company
Cochrane (1975) “An economic basis for the Gravity Model” JTEP Jan 1975
Cochrane (1976) “A more general form of the gravity model” Traffic Engineering and Control, 1976
Koenig 1974 “Theorie economique de l’accessibilite urbaine” Revue Economique
De Rus “The economic effects of high speed rail investment” OECD DP 2008-16, 2008
MVA 2007 “Rail passenger demand forecasting research – Workstream 1
Whelan & Lowe “Understanding the demand for rail travel” Transport Modelling Forum June 2009
Hausman 1997 “Valuation of New Goods under perfect and imperfect competition” In “The economics of new goods” Bresnehan (ed)
Evans1969 “Intercity travel and the London – Midland Electrification” JTEP
Rohr, Daly, Patruni & Tsang (2009) “The importance of frequency and destination choice effects in long distance travel behaviour: what choice models can tell us” ICMC, Harrogate
Greene (2008) “Econometric Analysis” 6th Edition, Prentice Hall