DfT LGV Forecasting Model: Overview of Delivered Model

Ref	**Location**
TN003	T:\10110012 DFT2_RD_FWK\H Project Master Files\07 TASK & variation orders\1_PRE-AWARD_AA\70004128_DFT2_332_LGV Modelling\Documents\TechNote\TN003 - Overview of Delivered Model.docx

To:	Liz Jacobs, DfT	**From:**	Tim Gent, WSP
Date:	30 June 2014	**cc:**	Ian Williams, Malcom Jay, Philip Sumner
Project Code:	70004128	**Ref:**	70004128-TN003

Version Control

Version	Date	By	Reviewed by	Authorised (TN only)	Notes
1.0	24/7/2014	Gui Gui	TJG		Initial Draft of note based on TN001
1.1	28/7/2014	TJG			Revised and extended as TN003. Draft Issue to DfT.
1.2	30/7/2014	TJG		***TJG***	
2.0	18/11/2014	QN	TJG		Comments added on updating directories. Section 3 on operation of the spreadsheet added.
2.1	18/12/14	TJG		***TJG***	**Reviewed, extended and authorised for release.**

1 Introduction

1.1 Purpose of note

1.1.1 This technical note outlines the revised the LGV model as a practical tool for use for forecasting by the Department. It provides detailed information on the derivation and application of the current spreadsheet model, including the enhancements and updates to the model that have emerged from the re-estimation of parameters and of model form that has been carried out in this study.

1.1.2 This new model has been designed to avoid unnecessary changes in approach to the original LGV model implemented and in use within the spreadsheet entitled LGV Model 2013.xls. Instead it extends the original model:

1) by introducing some further explanatory variables to take account of historic differences in traffic growth rates by region and road type;
2) through some other adjustments to the explanatory variables; and
3) through the use of more recent time series data for the estimation of parameters.

1.1.3 Further information on the formulation and interpretation of the updated LGV model and the sequence of tests used in its development, are covered in PN004 LGV_Model Design.docx. The data that is used to support the development of the model and its use for forecasting is described in TN002 Data for

Model.

1.2 Structure of Note

1.2.1 This note is structured as follows:

- Section 2: Overview of Forecasting Model. Main structure of the model as delivered, and discussion of model properties and the spreadsheet provided.
- Section 3: The Forecasting Model spreadsheet
- Section 4: Discussion of the Forecasts produced.
- Appendices A-C: Charts showing the model fit in the estimation period, and forecasts by region and road type.

2 Overview of Forecasting Model

2.1 Model Form and Equation

2.1.1 The original model that was estimated by Bradburn and Hyman (2002) had the following form

$$L_t = \beta_0 (L_{t-1})^{\beta_1} (L_{t-2})^{\beta_2} (F_t)^{\beta_3} (G_t)^{\beta_4} \tag{1}$$

The proposed model extends this formulation in a number of ways as shown in equation (2):

$$L_{rst} = \beta_{0rs} (L_{rst-1})^{\beta_1} (L_{rst-2})^{\beta_2} (F_t)^{\beta_3} \prod_i (G_{rt})^{\beta_{4i}} \tag{2}$$

where

L_{rst}	denotes the LGV annual traffic kilometres in region: *r*, on road type: *s*, in the year: *t*
F_t	denotes the average fuel price
G_{rt}	denotes the GVA in region: *r,* in the year: *t*
$\beta_{0rs}\ \beta_1 \beta_2\ \beta_3\ \beta_{4i}$	are estimated model parameters.

2.1.2 This model form was selected after a series of tests with different configurations. The current form corresponds to our Test 7 (see references below). The main features of this equation to note are as follows:

1) The forecast of the LGV traffic growth rate is differentiated by road type by region, rather than being a single growth rate common to all roads types and areas;

2) To facilitate this, the full set of constants β_{0rs} are used, denoting the different traffic growth rates observed in the range of combinations of road type and region;

3) First and second order LGV traffic lag terms are included by region and road type. As currently implemented, it is these lagged values and the β_{0rs} multiplier which provide the differentiation by region and road type;

4) Whilst GVA both nationally and by region were tested, it was found that national GDP was the most effective economic variable. The current model form operates with a single national GDP forecast;

5) The fuel price term was lagged by one year in some tests, because this has tended to perform a little better for some models in the estimation testing. However, it was found that an unlagged form performed better, so the lag has been removed in the final form. Unlike in the original 2002 model, the fuel price term is now statistically highly significant in our current tests;

6) The fuel price term used in the estimation is now a weighted average of the petrol and diesel prices, appropriately adjusted to discount for recoverable VAT payments and for reduced fuel use due to engine efficiency improvements. The original model had been estimated based on the diesel price series alone, though the model was later converted when forecasting to use this weighted average discounted price;

7) The inputs to the model have also been updated to provide a starting year for the model of 2013 and measurements of real prices in pounds.

2.2 Model Parameter Values

2.2.1 The Annex at the end of this note provides a full list of the alternative forms which were tested. The best results were achieved for Test 7 by using PLM model. The coefficient for test 7 compared with the original function and other best-performing tests are listed in Table 2.1

Table 2.1 Estimated coefficient values for the original and current models

Parameter	Variable	Original	Test 9	Test 7	Test 7iii
No. years		45	14	19	19
No. segments		1	55	55	55
$\ln(\beta_{0rs})$**	Intercept	– 3.585	-1.061	0.166	0.156
β_1	1st order lag traffic	1.096	0.797	0.964	0.924
β_2	2nd order lag traffic	– 0.386	0.042	-0.138	-0.098
β_3	Diesel price	– 0.073			
	Average fuel price (exc. VAT			-0.103	
	Lagged average fuel price (exc. VAT)		-0.083		-0.097
β_{41}	GDP	0.369		0.270	0.264
	Regional GVA		0.210		
β_{42}	Regional GVA – F		0.029		
β_{43}	Regional GVA – G		0.028		
Adjusted R^2			0.860	0.905	0.905
Date of estimation		2001	4/7/14	4/7/14	4/7/14
Fuel elasticity		-0.25	-0.52	-0.59	-0.56
GVA/GDP elasticity		1.27	1.66	1.55	1.52

*Source: Fixed effects model using standard panel estimation function *plm().*
*** Figures shown are estimated intercepts from the random model, whereas the forecast uses different values by segment. See section* ***Error! Reference source not found.*** *for a fuller description of the derivation of this coefficient by segment.*
Values shaded are not statistically significant at the 5% level.

2.2.2 Other than for the second order traffic lag, which now is a much smaller positive or negative value rather than a large positive value, the coefficients for the explanatory variables have retained the same sign and broad order of magnitude as previously in all three new model formulations. The relative magnitude of the coefficient for the intercept is not directly comparable between the old and new models due to the introduction of fixed effects to represent the region and road type segmentation.

2.2.3 The comparison of the elasticity values indicates that in all three new tests the fuel price elasticity of -0.59 is over double its 2002 value of -0.25. In a similar vein the new GVA/GDP elasticity 1.55 is considerably larger than the 1.27 found in 2002.

2.3 Data Sources for Forecasting

2.3.1 The source data used for forecasting is that contained in the spreadsheet LGV model 2013.xls that was supplied to us by DfT. It contains all the components that have been used to create the future year fuel price series.

2.3.2 The weighted average discount fuel price terms have been scaled so that they are in the price units of the recent LGV model 2013.xls database. The series in use is defined from consistent ingredients and conversions to real prices from our start date of 1993 through to 2013 and forward to 2040.

2.4 Saturation Effects

2.4.1 The previous DfT forecasts have included a factor to ‘cap’ growth in future years, based on the view that some saturation can be expected to occur. Saturation might arise either through demand side effects (e.g. because GDP growth ceases to be linked to a requirement for LGV kms to grow at the elasticity currently observed), or supply side effects (e.g. because road capacity is such that congestion effects make continued traffic growth relatively less attractive).

2.4.2 These are not effects that can be explicitly dealt with by the current model form, and there is no existing research or mechanism to suggest what level of saturation may in fact occur. Nevertheless it has been agreed with DfT that it is desirable for some levelling off of LGV growth to be built-in, as it seems reasonable to assume that this will occur in the longer-term.

2.4.3 The forecasts previously supplied (see TN001, 11 July 2014) exhibited reasonably constant percentage growth rates through the forecast period, giving rise to increasing absolute growth (i.e. no saturation effects). Our further experimentation with model estimations have not altered this basic position, and therefore we have considered possible mechanisms to introduce saturation artificially.

2.4.4 Two possible approaches were considered. First of all, we considered the view that ‘supply-side’ saturation effects are most likely to occur first in London and the South East, and on urban roads, the potential was considered to apply parameters from these regions and road types to other segments over time, but in a selective way. For example, urban areas of the North West and West Midlands might be expected to become ‘more like’ those of London, though this effect might not reach other regions until later. Similarly, some existing roads classified as rural may over time adopt urban characteristics.

2.4.5 However, this approach was abandoned as it was found that the substitution of parameters did not give intuitively sensible results. From observing the growth rates for 2003-2013, and the estimated β_{0rs} parameters, it was concluded that the model region/road type estimation is not yet showing saturation in ‘intuitive’ ways, and hence there is no basis currently to apply this approach.

Adjustment of GDP Elasticity

2.4.6 As an alternative, we have considered whether the GDP elasticity of the model can be lowered over time. This is the approach previously applied by the DfT in the 2010 model. To provide an improved evidence base for this measure, the long-term GDP elasticity of LGV km for Great Britain as a whole was studied. Figure 2.1 shows this trend; the elasticity has been calculated by dividing the percentage growth in LGV km by the percentage growth in GDP. This figure is shown both on an annual and 5 year basis (e.g. calculating growth between consecutive years, and 5 years gaps such as 1975 to 1980). Figure 2.2 and Figure 2.3 show the percentage growth rates corresponding with these for background.

2.4.7 This demonstrates a great deal of variation in the elasticity over the period. Elasticity was extremely high in the 1950s and early 1960s, though this perhaps relates strongly the very low usage of road transport for freight before this period, and can probably be discounted. From the late 1960s, the single year elasticity varies widely, though evidently some peaks and troughs are related to oil crisis of the 1970s, recession of the late 1970s/early 1980s and 1980s boom. Through these periods the 5 year figure remains comparatively stable, though varying above/below 1. The only exception is the early 1990s, where the 5 year elasticity reaches nearly 3, whilst the single year elasticity falls. The 5 year trend appears to be related to a large rise in LGV km in the second half of the 1980s and early 1990, whilst the single year fall is clearly related to the early 1990s recession.

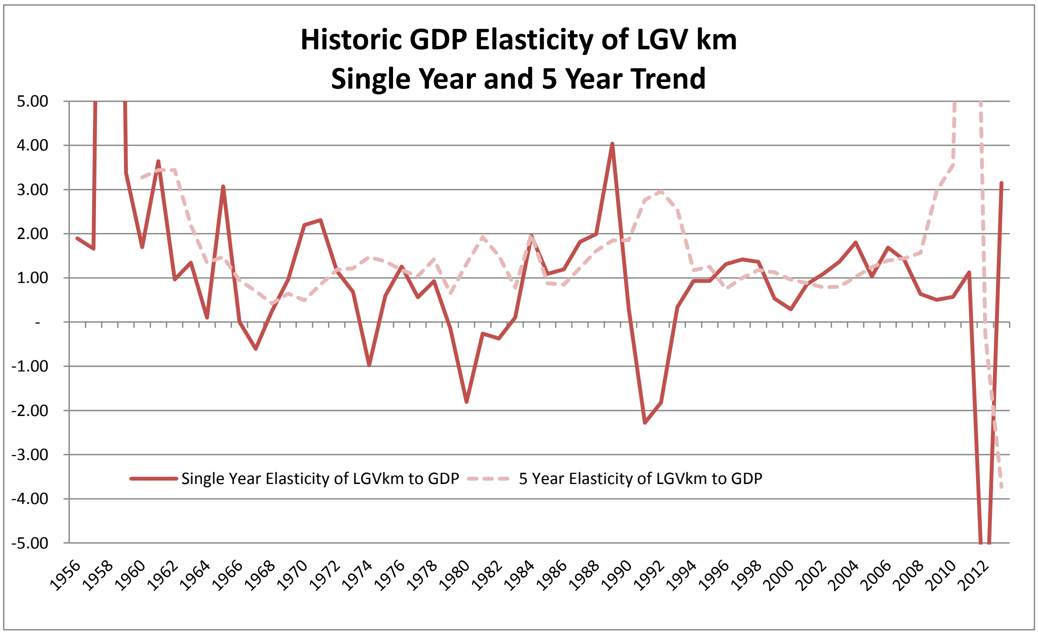

Figure 2.1 - Comparison of Single Year and 5 year elasticity of LGV km to GDP

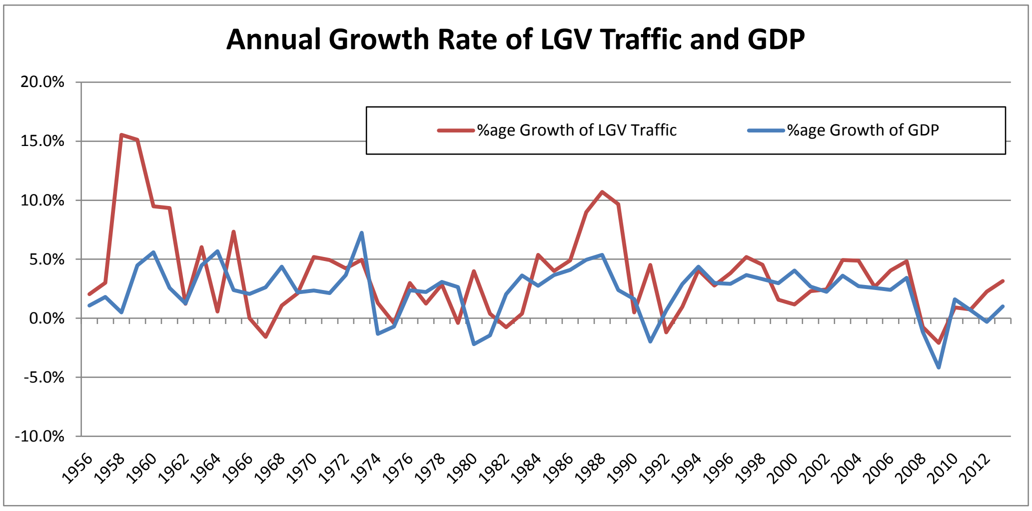

Figure 2.2 – Annual growth rate of GDP and LGV Traffic compared

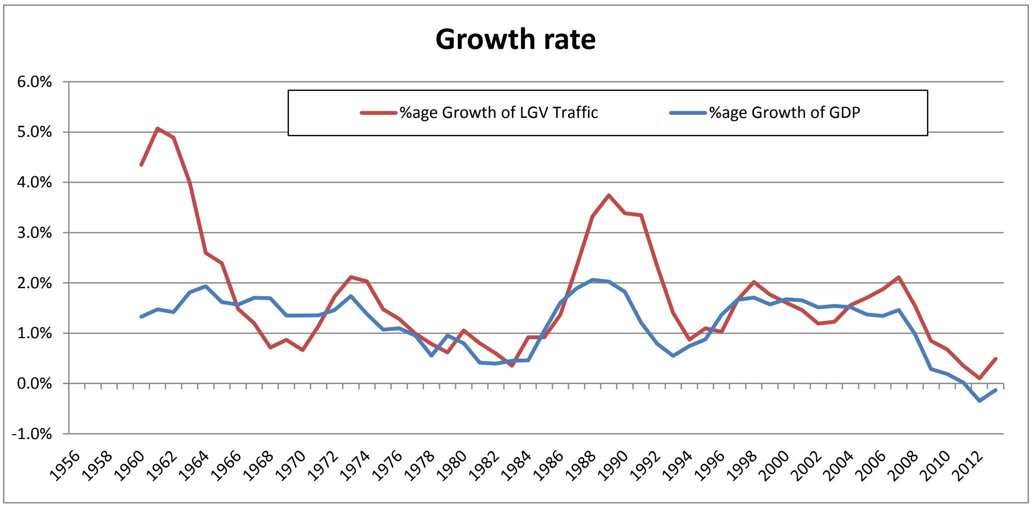

Figure 2.3 – 5 Year growth rate of GDP and LGV Traffic compared

2.4.8 Following the 1990s recession, both elasticities then vary around 1 until c.2005. At this point the 5 year elasticity starts to rise. Both annual and 5 year elasticities are then extremely erratic from 2010, but moving in opposite directions. It appears that two effects are occurring here:

- The LGV traffic grows at a rate above GDP growth for a sustained period through the mid-2000s, causing the 5 year elasticity to start to rise;
- At the onset of the recession, there was a fall in LGV traffic year on year, but there was not at any point a fall in LGV traffic over 5 years. This caused a very high 5 year GDP elasticity during this period (the negative value in 2013 corresponding to a point when GDP had fallen over 5 years, but LGV traffic had not).

2.4.9 Taking these observations together, it may be reasonable conclude that the long-run elasticity of LGV traffic to GDP varies equally above/below 1, apart from 3 periods in the late 1950s/early 1960s, late 1980/early 1990s and through the mid-2000s. During these periods LGV traffic growth out-strips GDP growth considerably, however in each case this occurred for approximately half a decade. Further research and consideration may be needed to identify causes for each of these periods. It is possible that each is related to a period when LGV use increased rapidly due to change in industry and haulage practices, technology available or structural economic drivers (costs and market changes).

2.4.10 It is not clear whether we are currently experiencing a similar growth: certainly LGV growth continued to outstrip GDP growth for most of the recession, and did so in 2013 compared with 2012. There is no evidence available from which we can conclude whether this will continue.

Application to the Forecasting Model

2.4.11 As shown in Table 2.1, the current model parameters result in a GDP-elasticity for LGV traffic of 1.55, compared with the 1.27 value found by the DfT in 2002. This does not appear surprising given that our estimation covers the period 1993-2013 during which elasticities have frequently been well above 1.

2.4.12 Based on the observations above, it appears reasonable to assume that GDP elasticity will fall back in future towards an average of 1. Based on the estimation and model form, we have concluded that 1 is the lowest elasticity which can be justifiably adopted within the model.

2.4.13 We have therefore assumed that the GDP-elasticity will fall uniformly from 1.55 to 1 through the period 2013 to 2040[1]. The spreadsheet has however been constructed so as to allow this assumption to be changed, with the GDP-elasticity altered every 5 years.

2.5 Model Start Year

2.5.1 By default, all forecasts were carried out using 2014 as the first modelled year, and with the lagged variables in that year based on the observed values for 2012 and 2013 respectively.

2.5.2 However, it was found that for Scotland this resulted in anomalous motorway growth in the years 2014-2016, with traffic at first falling. It was found that this is due to very high growth in motorway traffic in Scotland between 2011 and 2013 (11% in 2011-12 and 6% in 2012-13). This appears to relate to the increase in motorway length in 2012: the total motorway length in Scotland increased from 410km to 460km (*Source: DfT Statistics*), a growth of 11%[2].

2.5.3 It was felt that this rise in motorway length was not representative, and causing the anomalous behaviour of the initial motorway forecasts. We have therefore made the decision to set the model start year for Scotland to 2012 rather than 2014. This avoids prevents the 2011-12 motorway growth from affecting the lagged variables and having an undue influence on the forecast growth for Scotland.

2.5.4 It is possible that a further review of the change in road lengths during the model estimation period (1993-2013) could also be beneficial, though an initial check shows that growth is generally small in percentage terms.

[1] A target GDP is achieved in the model by altering the $\beta_4$parameter according to the equation: $GDP = \frac{\beta_4}{(1-\beta_1-\beta_2)}$

[2] We believe this growth is due to work on the M8, M74 and M73, much of which upgraded existing roads.

2.6 Properties of the model form

2.6.1 There are two properties of the model which we have observed during experimentation with the forecasts, and are worth noting here. These may be of interest in considering the forecasts, however we do not believe that either give any cause for concern.

2.6.2 Both of these arise from the use of the lagged variables: study of equation 2 and the parameters in Table 2.1 will show that the lagged variables parameters are of opposite signs. This has the effect that the 1 year lagged LGVkm is divided by the 2 year lagged LGVkm (at a lesser power). This has the following long and short-term effects:

- **Short-term:** If the lagged data rises very rapidly between the 2^{nd} and 1^{st} lag year, then this can produce a multiplier <1, causing the forecast LGVkm to fall. This is what occurred in 2014 for Scottish Motorways, as described in paragraph 2.5.2 above. The opposite will occur for sharp falls in the lagged data. This may be viewed as a 'damping' of sudden changes in LGV km. However, this is a rare effect and is only likely to come into play in the years where observed data is used.
- **Long-term:** The ratio overtime becomes a nested series of ratios of the forecast values. This tends towards 1 as the number of forecast years increases, as both the lagged variables and the β_{0rs} variables cancel out. This means that in later years the model growth rate is dominated by the un-segmented variables related to GDP and fuel cost, rather than the segmented variables.

2.6.3 The one case where we believed the short-term effect to be an issue has been removed from the model. The longer-term effect we believe may be an appropriate property of the model: it ensures that the growth is differentiated by region and road type in the early forecast years, but that the influence of this declines over time, when extrapolation of these differences becomes less relevant.

3 The Forecasting Model Spreadsheet

3.1 Introduction

3.1.1 The LGV Forecasting spreadsheet provided (currently version 4.02) has been constructed to allow the forecasts discussed in this report to be produced and viewed. It also allows alteration of key inputs and assumptions to create new forecasts in a straightforward and robust manner.

3.1.2 This section provides a brief overview of the spreadsheet, the main calculations and steps required to create new forecasts. For further advice on the operation, update and interpretation of the spreadsheet please contact WSP.

3.2 Overview of Worksheets

3.2.1 The spreadsheet model consists of 21 worksheets:

- The first worksheet <<ReadMe>> gives a brief introduction and explanation of the whole workbook and the evolution of the version changes.
- <<LGV forecasting>> (coloured in purple) constructs the forecasting model with inputs coming from tabs <<Vlookup>>, <<ParameterVals>> and <<Fuel and GDP forecast>> (coloured in light yellow).
- The red tabs provide charts of the model results, which are created from the data processing tabs <<PivotnChart_LGVIndex_region>>, <<PivotnChart_LGVkm_region>>, <<Comparison_Abs>> and <<Comparison_Index>> (coloured in light purple) serving different analytical purposes.
- The tab << Comparison of elasticity>> compares the elasticity values created by this forecasting model against those of DfT's earlier forecasting model; while tab << Elasticity of Past trend>> studies the pattern of elasticity over the past decades.
- << COBA>>, <<NATCOP>>, <<GDP & Fuel DFT>> and << DFT LGV Model>> are jointly used to produce output in tab <<Fuel and GDP forecast>> and were supplied originally by DfT.

3.3 LGV Forecasting Sheet

3.3.1 The main forecasts are produced on the worksheet << LGV forecasting>>. To explain the functioning of the spreadsheet, this sheet is discussed in more detail in this section.

3.3.2 The sheet is in rows, with each row containing an LGV forecast for a specific combination of year (column F), region (column G) and road type (column H). See the <<Vlookup>> sheet for the meaning of the region and road type codes.

3.3.3 The **output LGV km forecasts** are in column W, expressed in billion vehicle km per year, and in column Y the values are indexed (the index year is set by the input value in column AA).

3.3.4 This sheet format is useful for calculation, but not ideal for viewing model results. The results are best analysed using pivottables which allow the forecasts to be tabulated and filtered by region or road type. The pivot sheets provide this presentation, and supply data for the charts within the spreadsheet.

3.3.5 **Parameter values:** Columns I, J, K, O and Q import the β values calculated in tab <<ParameterVals>> for the correct year.

3.3.6 **Fuel price and GDP data** (columns P and R) from 1993 to 2013 are imported from <<Fuel and GDP Forecast>>, the source of this data ultimately being the DfT's original LGV forecast model.

3.3.7 **Observed Traffic data for the given year, region and road type** is stored in column L is the traffic data from year 1993 to 2013 in billion vehicle km taken from column Q (TRLGV) in tab <<LARoadTraffic19932013>> of LARoadTraffic19932013_processed_V6.xlsx, which were originally provided by DfT.

3.3.8 Columns M and N select the correct **lagged LGV km variable**, selecting either observed or modelled values for the lag. Observed LGV km are used where these exist and are reliable, as this ensures the forecasts for the near future are as accurate as possible. The precise year at which forecast data is used varies by region, based on values in the <<Vlookup>> sheet. provide is calculated from the surveyed traffic data (column L), which should be automatically updated if column L is changed.

3.3.9 Column P is the **input fuel price**, imported from the sheet <<DFT LGV Model>>[3] column N.

3.3.10 Column R is the **GDP input**, from the tab <<Fuel and GDP forecast>> column D (RPI GDP)[4].

3.4 Altering LGV Forecasts

3.4.1 Major stages of the LGV forecast calculation are as follows:

- **Forecasted LGV km and index:** The forecast LGVkm is output in << LGV forecasting>> in column W (absolute term) and column Y (index term), but can be viewed in more detail in the pivot sheets and charts.
- **Base year for indexing:** The spreadsheet allows the user to select which year is used as base for indexing in the charts, as this can vary the growth presented and is worth experimentation. To allow flexibility, we have this mechanism in the tab << LGV forecasting>> in cell AA5. Simply select the target year from the drop-down list and then click the button "Refresh Pivot".
- **Parameter changes:** All the parameter values used in the forecasting model equation are stored in <<ParameterVals>>. The original values are copied from the output of R-statistics test. β_0 values are the exponential term of the original intercept values from the output and varied by region and road. The sheet allows a target GDP elasticity to be set, and the β_4 parameter to then decrease through time to accommodate this, as described in section 2.4 above. Currently, the target elasticity value is set to 1 in 2040. To revise this, simply change the value in cell V69 in this tab and then click the button "Refresh Pivot" on the right to make sure all the pivot tables and charts are updated using the new calculated values. Though other parameter values can be changed through time, we do not currently advise this, and no facility for 'smooth' changes through time is provided.
- **Updates to Fuel and GDP Inputs:** The user is able to make changes to input values for GDP and fuel prices in the sheet <<Fuel and GDP Forecast>>. It is advised that any changes are entered directly into columns C and D, either over-writing current formulae, or re-directing the references to a new input sheet. Both actual historic figures and revised forecasts (perhaps for generating scenarios) can be input here directly as a single column. It is suggested that values are clearly annotated to show which are observed and which forecast.
- **Updates to Observed LGV km:** As further data on observed LGV kilometres becomes available each year, the data in column L of <<LGV Forecasting>> can be updated, or further rows of data added. To ensure that the data is used in the forecast, it is necessary to update the table in row 18 of <<Vlookup>>, increasing the year at which modelled values are used for the lagged variable.

3.4.2 After any change, the values in column W of <<LGV Forecasting>> should update at the next re-calculation. However, the charts and tables in other sheets require that the pivottables are updated, which is achieved by pressing the 'Refresh Pivots' button in column AB of <<LGV Forecasting>>

[3] <<DFT LGV Model>> is based on the LGV model 2013 provided by DfT. It provides the Fuel price from year 2003 to 2040. The fuel price was extended backwards from 2003 through to 1993 in a form that is maximally consistent with the corresponding data already there from COBA from 2003 through to an extrapolated 2040. The method of extension was described in the spreadsheet in detail.Most of the data in the tab <<DFT LGV Model>> is from the tab <<COBA>>, which was provided by DFT. The columns G, V, H, W, AT, AU, BB, BL, and CI in the tab <<COBA>> can be updated if new data about the Fuel price is available. Then the average Fuel price used in forecast model should be automatically updated.

[4] From year 1993 to 2010, the GDP value used in the forecast model is from tab <<GDP & Fuel DFT> > column D (RPI GDP), after 2010, the value is from forecasts in tab <<NATCOP>>>. The content in tab <<GDP & Fuel DFT>> is the same as the content in the tab <<Data and Regression Analysis>> in the spreadsheet 130501 Data for LGV Regression - Re-estimation.xls provided by DFT. Column D (RPI GDP) in the tab <<GDP & Fuel DFT>> can be updated and the tab << Fuel and GDP forecast>> should be updated automatically. In the Tab << NATCOP>>, column E (GDP) is used from 2011 to 2040 and can be updated and the tab << Fuel and GDP forecast>> should be updated automatically.

4 Discussion of Forecasts

4.1 Introduction

4.1.1 This Section briefly analyses the results produced by the most recent version of the forecasting model. These are from the New LGV Forecasting Model v4.nn, which is based on model test 7. The reduction in GDP elasticity from 1.55 in 2013 to 1.00 in 2040 has been applied, as discussed in section 2.4.13 above.

4.2 Model Results

Discussion of Modelled vs Input values

4.2.1 Figures 3.1 and 3.2 below show the modelled LGV kilometres vs the input data for the historic data period, both by road type for all of GB, and by region for each road type. Appendix A contains a full set of charts by region and road type. Note that these forecasts are specifically prepared using the modelled LGV km values as lagged variables during this period NOT the observed values. This gives a reasonable impression of the model performance over the model fit period.

4.2.2 Both the charts shown here and those in Appendix A demonstrate that the modelled values are tracking the input data well across the historic period. The relative growth by region and road type is well captured, including for example the much higher growth rate in the South East region and lower in the North East, and the relatively low growth on Urban compared with Rural A Roads.

4.2.3 There may however be some suggestion that the modelled values are not responding to observed trend from 2010 to 2013: Motorway and Rural A road volumes have risen in that period, whilst the modelled figures remain relatively flat. Similarly, some recent growth trend in the regions (particularly the South East and North West) is perhaps not as evident as might be desired. Study of the charts in Appendix A shows that this is related solely to Motorway growth in the North West, but across both Motorways and Urban A roads in the South East (with Rural Other roads showing a counter-trend).

4.2.4 Otherwise, the Appendix A charts again show that the model is representing the relative growth rates reasonably well across this period. The only other point of note is the relatively large variation in the INPUT data for Urban A roads in London.

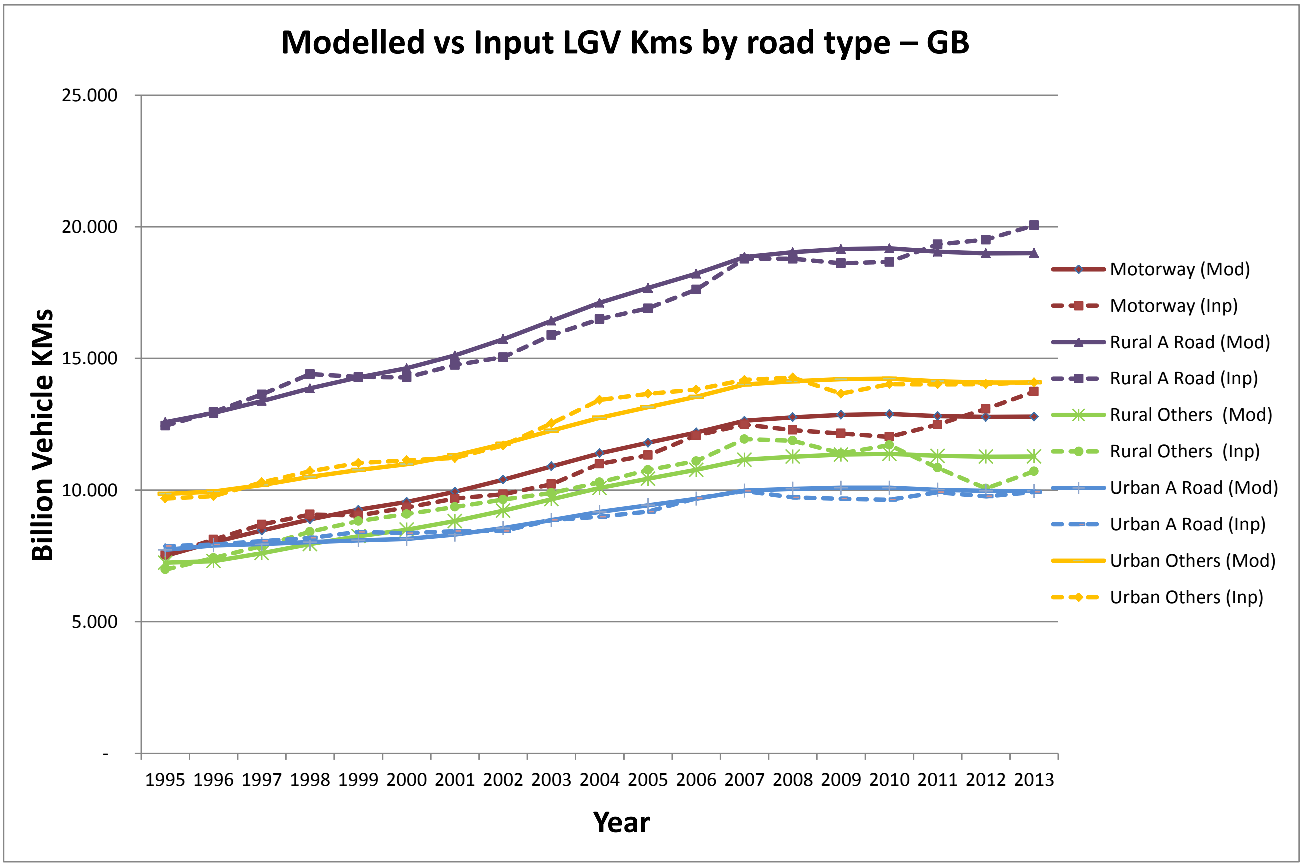

Figure 4.1 - Model Fit by Road Type (National) - Input vs Modelled LGV kms

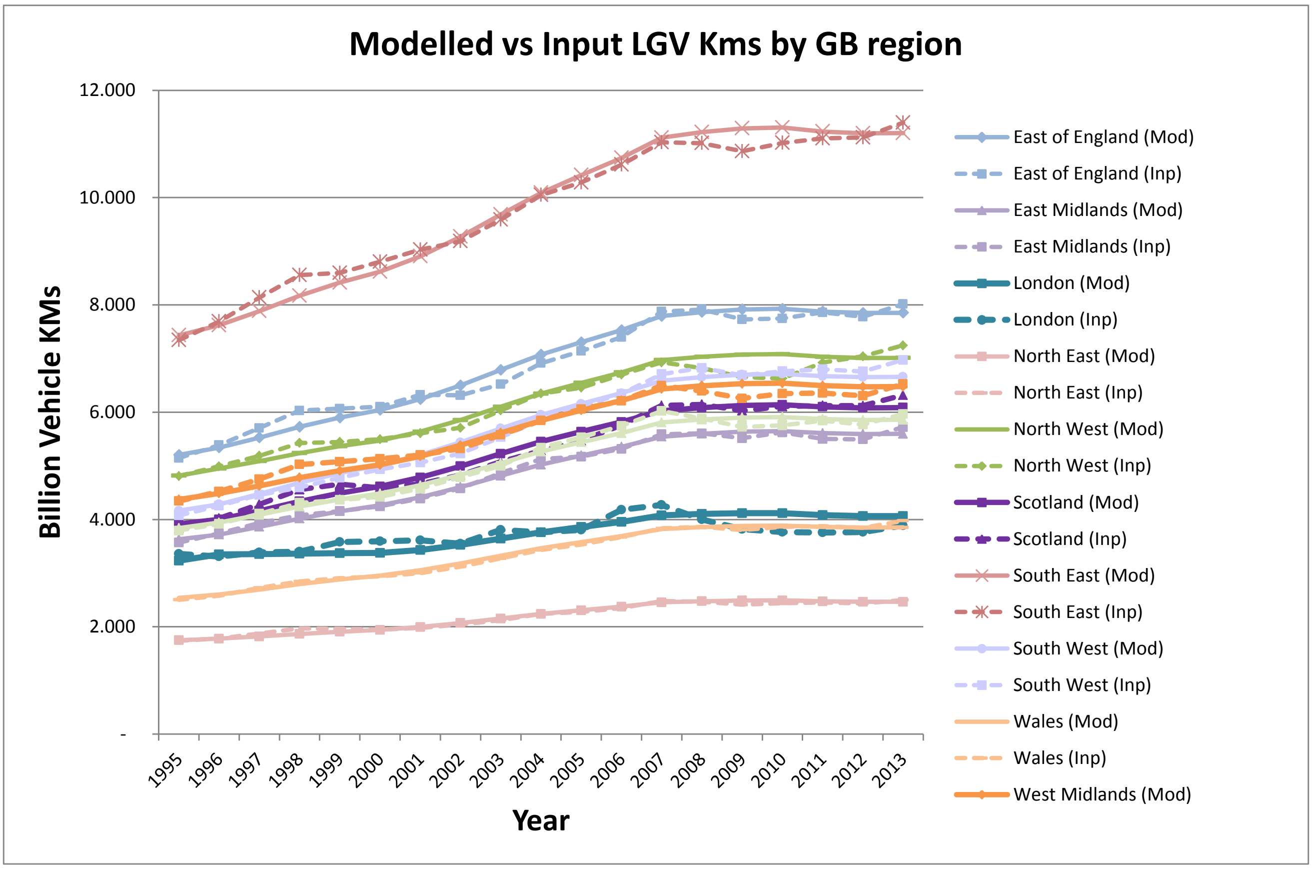

Figure 4.2 - Model Fit by Region (All Road Types) - Input vs Modelled LGV kms

Forecast Model Results

4.2.5 Appendix B contains a full set of charts showing the modelled figures both for the historic period (shaded) and the forecast years. For ease of reference, these charts are indexed against the 2013 value, and each chart includes for comparison the GDP and Fuel Cost input variables also indexed to 2013. Appendix C provides charts with the absolute LGV km for each region, broken down by road type.

4.2.6 The indexed charts show that the growth rates are higher in the short-term, and growth declines through the forecast period. This is due to the decrease in GDP elasticity which has been implemented, as described in Section 2.4. In general, all regions and road types now show forecast growth of between 60% and 80% from 2013 to 2040. As Appendix C shows, this includes a very wide variety of absolute growth.

5 References

Bradburn P & G Hyman (2002) *An Econometric Investigation of Car Use in the National Transport Model for Great Britain*. Proceedings of the European Transport Conference, 2002.
Green W (2012) *Econometric Analysis, 7th Ed*. Prentice Hall.
Hansen B E (2014) *Econometrics*. University of Wisconsin, Department of Economics.
URL: http://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf

Annex – Notes on Model form and List of Model Estimation Tests in R

The constants β_{0rs} for use in the forecasting equation (2) above are generated as exp(a + b_{0rs}), where "b_{0rs}" are the segment specific coefficients and "a" is the intercept estimated by the log linear regression equation. In the R software, an overall intercept "a" is automatically output for each estimated model. Where there are factors (dummy variables) included in the explanatory variables, one category (by convention, the first in the factor list) is dropped and so the intercept corresponds to the coefficient for this segment. The constants for the other segments are calculated by summing the intercept and the segment specific coefficient in the exponential transformation shown above.

For a random effects model the set of coefficients will automatically be output for any factors explicitly included in the explanatory variable list, along with an intercept coefficient value. In a fixed effects model the full set of segment specific constants are all calculated automatically but the function *fixef()* must be called to output them for use. In its default output option "type='level'" these coefficients from the fixed effects model are output in a form in which the overall intercept term is zero. Their values match exactly to the corresponding random effects model provided that this has an explicit factor term for every individual segment. It is however necessary that the random effects intercept value is added back onto each random effects coefficient, as explained in the previous paragraph, for the match to be obvious.

The following tests have been implemented in the estimation process to find out the best model. Further explanation is given in PN004 V2.4 (*Technical Note forthcoming*).

Test Number	Formula for Estimation in R	Lagged Variable	Tests
5	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel) + log(GDP) + Region	LGVkm (1 and 2 years)	PLM/PGMM
Full Lagged 5	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GDP1) + Region	LGVkm(1 and 2 years), Fuel, GDP	PLM
Fuel Lagged 5	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GDP) + Region	LGVkm(1 and 2 years), Fuel	PLM
6	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel) + log(GDP) + RT	LGVkm(1 and 2 years)	PLM/PGMM
Full Lagged 6	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GDP1) + RT	LGVkm(1 and 2 years), Fuel, GDP	PLM
Fuel lagged 6	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GDP) + RT	LGVkm(1 and 2 years), Fuel	PLM
67	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel) + log(GDP) + Region +RT	LGVkm(1 and 2 years)	PLM/PGMM
Full Lagged 67	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel`) + log(GDP`) + Region +RT	LGVkm(1 and 2 years), Fuel, GDP	PLM
Fuel Lagged 67	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel`) + log(GDP) + Region +RT	LGVkm(1 and 2 years), Fuel	PLM

Test Number	Formula for Estimation in R	Lagged Variable	Tests
7	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel) + log(GDP) + RnRT	LGVkm(1 and 2 years)	PLM/PGMM
7i	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GDP1) + RnRT	LGVkm(1 and 2 years), Fuel, GDP	PLM
7iii	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GDP) + RnRT	LGVkm(1 and 2 years), Fuel	PLM
8	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel) + log(GVAAll) + log(GVAF) + log(GVAG) +Region	LGVkm(1 and 2 years)	PLM/PGMM
8 Full Lagged	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GVAAll1) + log(GVAF1) + log(GVAG1) +Region	LGVkm(1 and 2 years), Fuel GVAAll, GVAF, GVG	
9	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel) + log(GVAAll) + log(GVAF1) + log(GVAG1) +RnRT	LGVkm(1 and 2 years)	PLM/PGMM
9 Full Lagged	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GVAAll1) + log(GVAF1) + log(GVAG1) +RnRT	LGVkm(1 and 2 years), Fuel GVAAll, GVAF, GVG	PLM
9 Fuel lagged	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GVAAll) + log(GVAF) + log(GVAG) +RnRT	LGVkm(1 and 2 years), Fuel	PLM
9 Fuel, GVA,GVG lagged	log(LGVkm) ~ log(LGVkm1) + log(LGVkm2) + log(Fuel1) + log(GVAAll) + log(GVAF1) + log(GVAG1) +RnRT	LGVkm(1 and 2 years), Fuel GVAF, GVG	PLM
9i	log(LGVkm) ~ log(LGVkm1) +log(Fuel) + log(GVAAll) +RnRT	LGVkm(1 and 2 years)	PLM
9ii	log(LGVkm) ~ log(LGVkm1) +log(Fuel1) + log(GVAAll)	LGVkm(1 and 2 years), Fuel	PLM