Variable Demand Modelling – Key Processes

TAG Unit 3.10.3

October 2009

Department for Transport

Transport Analysis Guidance (TAG)

This Unit is part of a family which can be accessed at www.dft.gov.uk/webtag

Contents

1 Variable Demand Modelling – Key Processes **1**

1.1 Background 1
1.2 Elasticity Methods or Full Variable Demand Modelling 2
1.3 Functional Form and Overview of Variable Demand Models 2
1.4 Which Responses? 5
1.5 Form of Models 7
1.6 Which Modes at What Detail? 11
1.7 Trip Distribution 14
1.8 Time of Day Choice 18
1.9 Hierarchy of Responses and Composite Costs 19
1.10 Local Calibration of Demand Models 23
1.11 Illustrative Parameter Values 26

2 Further Information **29**

3 References **30**

4 Document Provenance **30**

1 Appendix 1 Elasticity Models **31**

1.1 Functional Forms and Parameter Values of Elasticity Models 31

2 Appendix 2 Functional Forms for VDM **34**

2.1 Detailed Advice on Functional Forms of VDM 34

3 Appendix 3 Example Estimation of Modal Split **37**

4 Appendix 4 Incremental Model Formulation **39**

5 Appendix 5 Absolute Model Formulation **43**

1 Variable Demand Modelling – Key Processes

1.1 Background

1.1.1 TAG units 2.9 and 3.10 explain why variable demand modelling needs to be considered and provide guidance on how to carry out such modelling for highway schemes. This unit forms stage 3 of the variable demand process and provides the detailed advice required for those carrying out variable demand modelling. TAG Units 3.10.1 and 3.10.2 detail previous stages with TAG Unit 3.10.4 forming the subsequent stage.

1.1.2 While the guidance is based on cases where a full multi-modal model with appropriate segmentation and representation of demand responses is required, there are various places in WebTAG where it is noted that less detail may be acceptable, though it will usually be expected that an appropriate case is made for any simplifications adopted.. In particular, readers should consult the section 1.4 in Unit 3.10.1 regarding the need for a full representation of alternative modes.

1.1.3 In the case of schemes with a capital cost below £20 million, additional simplifications in the treatment of the demand model are acceptable (subject to the same justification), and these are also described in Unit 3.10.1, Section 1.5. In this particular case, paragraph 1.2.3 of this Unit does not apply.

1.1.4 In all other cases, the Guidance set out in this Unit will apply. Important recommendations are shown highlighted and in bold. If those actions are not followed analysts will need to provide rigorous justification for the course of action taken.

1.1.5 The key summary of this part of the advice is as follows:-

- Most variable demand models use some form of "**hierarchical logit" formulation**, in which the choice between travel alternatives (frequency, modes, destinations, time periods) depends upon an exponential function of the generalised cost or disutility.

- It is expected that distribution models will be included in all variable demand models. Details of the different model formulations are discussed, as is the representation of the fringes of a study area, which is particularly important when using trip distribution models.

- The representation of different modes in the variable demand model is discussed, and how it is normally acceptable to include the alternative mode(s) merely as a set of fixed costs. Occasionally, however it may be necessary to model the journey components in detail, including the effect of changing road conditions on bus travel.

- The modelling of departure time choice as a demand response or in close association with assignment is discussed. It is recommended that large "macro" adjustments only need to be modelled when considering differential pricing between time periods, or access restrictions

- Wherever possible the variable demand mechanisms should be calibrated on local data, to reflect the local strengths of the choice mechanisms. This is not always possible, however illustrative values obtained from a review of UK transport models are reported, which may be used when accompanied by realism tests where it is deemed too difficult to establish local values.

1.2 Elasticity Methods or Full Variable Demand Modelling

1.2.1 **"Own-cost" elasticity models** assume that the demand for travel between two points is purely a function of the change in costs on that mode between the two places. The strength of that function can vary for different trip lengths.

1.2.2 However, a own cost elasticity model applied to all trips cannot recreate all of the changes in trip lengths which are forecast by a trip distribution model, nor can it properly represent the transfer of trips from one mode to another or from one time period to another when there are changes to the cost of several modes. The importance of these deficiencies may vary from study to study, and research has been undertaken to assess the importance of these theoretical deficiencies for scheme appraisal. The initial results of this research suggest that elasticity models may significantly overestimate the effect of variable demand responses on scheme benefits, giving an overestimate of reductions from fixed trip matrix appraisal benefits. It is anticipated that further research will be commissioned by the Department to better understand the deficiencies of elasticity models.

1.2.3 Pending further research it is recommended that own cost elasticity models are not used instead of full variable demand models.

1.2.4 However, whilst an elasticity approach should not be used as a direct replacement for a full variable demand model, there may be a role for elasticity models in option testing. Once a variable demand model has been set up, then an elasticity model may be developed to reflect the results of this in terms of scheme benefits, (adjusting elasticity values so that user benefits of a given scheme are the same using the elasticity model as when using the fully specified variable demand model), which can then be used to refine a large number of scheme options. This method may be attractive where the full variable demand model requires significant run times and testing a large number of options would be impractical.

1.2.5 If an elasticity approach is to be applied in this way, then Appendix 1 sets out the different possible formulations.

1.3 Functional Form and Overview of Variable Demand Models

1.3.1 Most variable demand models use some form of "hierarchical logit" formulation, in which the choice between travel alternatives (frequency, modes, destinations, time periods) depends upon an exponential function of the generalised cost or disutility. That mechanism, and its application in variable demand models, is described here.

1.3.2 Any model of the demand for travel relies on a mathematical mechanism which reflects how demand will:

- fall if the generalised cost (time or money) increases; or
- rise if the cost decreases.

For example, the "own cost" elasticity mechanisms described in Appendix 1 modify an earlier estimate of demand using a curve which is as illustrated in Figure 1: the demand approaches (and is asymptotic to) zero, but never actually falls to zero even at very high costs.

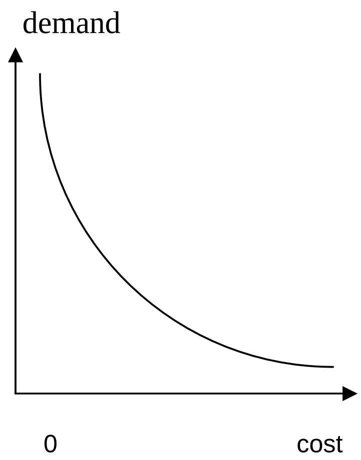


Figure 1 "Own cost" elasticity

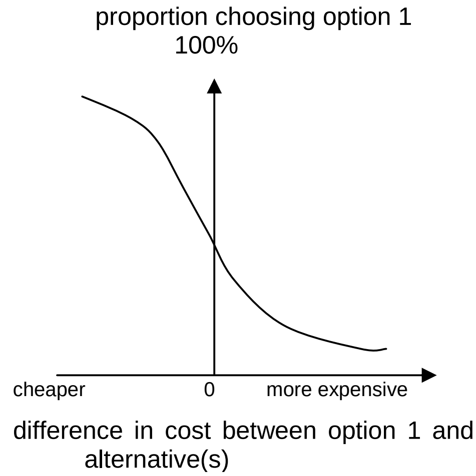


Figure 2 Choice model

1.3.3 In a variable demand model, a different mechanism is normally used to apportion the total demand in a particular travel category between two or more available choices, as for example between car and public transport, or between many different destinations. In this case, when the generalised cost of the specified choice is very much lower than the alternative choices almost all travellers will choose it, and if it is very much greater then very few will. Again, the function used is asymptotic to 100% or zero, as illustrated in Figure 2, since in any large population of travellers there will be a small but finite number who will take the apparently expensive choice. This arises even when only considering travellers with a car available and in general reflects how individual circumstances or choice preferences may be very different from the average.

1.3.4 The choice is unlikely to be based purely on simple formulations of generalised cost of travel: it will also depend upon appropriate zone or mode specific constants estimated in model calibration or implied by incremental models (see Section 1.5). They represent any extra utility gained by making that choice. For example, some modes will be inherently more attractive than others. Models assume that these hidden differences in the utility of travelling to any particular destination remain unchanged through time, and most therefore only need to reflect changes in the specific terms included in the generalised cost, as explained in *VDM Scope of the Model* (TAG Unit 3.10.2).

Logit Formulation

1.3.5 Appendix 2 describes some alternative forms of the many mathematical functions that can represent the behaviour shown in Figure 2. Logit is commonly used because it is easy to manipulate mathematically. In general, it is formulated as:

$$\mathbf{P_p = exp(-\lambda U_p)/\{\Sigma_q exp(-\lambda U_q)\}}$$

Where Pp is the proportion of travellers choosing alternative p out of q possibilities, Up is the disutility of option p (based on composite costs at lower levels in the hierarchy), and the summation in the denominator is over all q alternatives, including p. If there are only two choices this is called a binary logit model with the simple formulation $P1=exp(-\lambda U1)/[exp(-\lambda U1)+exp(-\lambda U2)]$; for more choices it is referred to as a multinomial logit model as described in more detail in Appendix 2.

1.3.6 The mechanism should be applied separately to different categories of travel, such as trip purposes, as the sensitivity of the model is likely to be different in each category. For different trip purposes, for example, the logit sensitivity parameter λ is likely to be numerically larger where there is more freedom to choose. **Thus more optional travel, such as shopping trips, tends to be more elastic and have a numerically larger λ value than, say, travel to work.**

Hierarchical Logit

1.3.7 A single logit model may be applied to the entire range of choices available using a ***multinomial*** (i.e. many choices) ***logit model***. However, that would implicitly assume that the sensitivities of those choices were all the same. Experience and intuition suggests that this is unlikely to be the case. This leads to a ***hierarchical*** system of logit formulations in which at each level a limited number of choices are considered. For example, a variable demand model might:

- first estimate the number of trips from any given origin (trip frequency - usually as an elasticity formulation);
- then estimate how many trips will choose each available mode (mode split); and
- then estimate how these trips choose amongst the available destinations (trip distribution)

(Note: this example excludes any time of day choice mechanism.)

1.3.8 The sequence chosen often varies between types of trip and does not necessarily represent the sequence of *thought* that makes these decisions. All the choices are interconnected, so that in a model that is converged, choices made earlier in the sequence are consistent with choices later in the sequence as the calculation is repeated. As a consequence the model has to be iterated to equilibrium.

1.3.9 Choices made higher in the hierarchy act as constraints on those made later. Hence, if the sensitivity of choice does not increase down the sequence there is a danger of later choices being too strongly influenced by earlier choices. Further discussion of the hierarchy of responses can be found in Section 1.9.

1.3.10 Within any of the steps (known as hierarchical levels), it may be desirable to model some secondary choices, for example, because travellers seem not to discriminate between different public transport modes in the same way as they treat the choice between car and public transport. Consequently, it may be preferable to split mode choice into a "high-level" two-way choice between car and public transport, with a "lower level" split into the different public transport modes, although this level of complexity is unlikely to be warranted for road scheme appraisal. This is often referred to as ***nested*** logit. It avoids the common problem with forecasting trips across three modes, car, bus and rail, of making the choice over-sensitive to changes in what travellers perceive as competing public transport services.

Application of the Nested Logit Function and Composite Costs

1.3.11 The logit function can be used in each stage of the variable demand model, in slightly different formulations, as described in Appendix 2. However, at each level in the hierarchy, it is necessary to calculate a disutility or generalised cost. For all except the lowest (last) level in the hierarchy this cost must reflect the available choices at lower levels: because it encompasses more than one type of choice it is called a *composite* cost or composite disutility. For example, the cost used in the trip frequency stage must represent a weighted average across the choices made in mode choice, where included across the time periods chosen, and across destination choices. **It is not sufficient to take a simple average across all the available choices, since the more costly choices will rarely be chosen. For this reason the "composite cost" incorporates the probability that a given choice will be made.** The logit formulation has the advantage that its composite cost is normally easy to calculate and is normally referred to as a "logsum" composite cost (see Section 1.9).

1.3.12 The logit mechanism can be applied in two ways:

- as an ***absolute*** prediction, to predict the absolute numbers of trips made between each origin and destination, by each mode, in each traveller category, and each time period;
- as an ***incremental*** model, which predicts the relative changes from a base case.

Hybrid models may make absolute calculations for some levels of the hierarchy, and apply incremental calculations to others. These aspects are discussed further in Section 1.5.

1.4 Which Responses?

1.4.1 Although this advice gives guidance on the four variable demand mechanisms – trip frequency, mode choice, trip distribution, and time of day choice – it may not be necessary to include all of them in your demand model. Which to include depends upon the circumstances and policy interests of your assessment, and also on the data you have available and the amount of effort which seems justified by the particular application.

1.4.2 Where available a higher level model may provide an indication of the relative importance of the mechanisms. Nevertheless, there is a wide consensus that trip frequency is less sensitive to changes in travel costs than either trip distribution or mode choice. Mode choice for those with a car available will have a strong effect only where public transport offers an acceptable alternative. Time of day choice is likely to be important if there is substantial traffic congestion in the area of interest, and if the forecast changes in demand increase or reduce this appreciably, or if some form of differential charging or access is to be examined.

Trip Generation and Frequency

1.4.3 It is important to understand that the trip generation/frequency stage involves two rather different aspects:

- modelling **trip generation** as a function of the demographic and socio-economic characteristics of the area; and
- implicitly representing the response of the trips so generated to changes in travel cost (this is referred to as **trip frequency**).

If the population or car ownership or built development of the area is changing appreciably over the time period of interest, then the first of these mechanisms will be essential. Once that growth has been established for the reference case (see *VDM Scope of the Model* **(**TAG Unit 3.10.2)) it may be appropriate to make forecast trip numbers responsive to travel costs. This requires the incorporation of the elasticity mechanism to represent trip frequency as described in Appendix 2.

1.4.4 If the modal split mechanism includes slow modes, then overall trip rates will be fairly stable and there will be no need to model the response of trip frequency to changes in travel cost. However, in this case the mode-choice modelling will have to be more complex. Even if slow modes are not included, but there is a realistic representation of choice between car and public transport, the effect of trip frequency is likely to be small. However, in general, the inclusion of a trip frequency mechanism requires no additional information beyond that required for trip distribution, and the only extra complication involved is the calculation of a mean composite cost of trips from each zone (see Section 1.9).

Mode Choice

1.4.5 The mode choice mechanism allocates trips to each of the main modes included in the model. The importance of this mechanism will depend upon how competitive other modes are with car travel on the road network affected by the scheme. It will also depend on the willingness of people to transfer between car and other modes. Generally, the relevant alternative mode is public transport, though if the scheme is used by short local trips slow modes may offer some alternative. Note that the modelled modal split should consider public transport usage by those with a car available for the journey. For this category of person the mode-choice will inevitably be much more inclined towards car use than the observed modal split overall. The importance of the mechanism can be judged by the prevailing modal split between car and the alternative for such journeys, as described in Section 1.6. Where modal split to public transport is deemed to be important, then mode-choice will need to be considered in the demand modelling. If the scheme is expected to have a great impact on public transport and/or public transport alternatives are to be tested it will be desirable to model the public transport options in greater detail (see Section 1.6).

1.4.6 A few models omit the mode choice mechanism altogether because modal transfer is not considered to be important. This is not recommended (see Section 1.6), but if that approach is used it will be important to include a trip frequency elasticity at a greater strength than usual, since this will act as proxy for trips transferred to the car mode from other modes and vice versa.

Trip Distribution

1.4.7 Trip distribution models spread the forecasts of generated trips over the available destinations, depending on the generalised cost of reaching that destination. This leads to an estimate of the number of trips between each pair of zones and intrazonal movements.

1.4.8 **When modelling individual variable demand responses it is expected that a distribution mechanism will be included**. This can have a substantial effect on the trip pattern and the amount of traffic using the scheme. Note that its effect on road traffic arises from its changing the mean length of trips over a number of years rather than the numbers of trips, but its net effect on demand between individual origins and destinations and on economic benefit is likely to be stronger than mode choice.

Time of Day Choice

1.4.9 It is unlikely that a variable demand model for a road-based scheme will need to look at time of day choice over all 24 hours in a day but there will be circumstances where the choice of time of travel in certain parts of the day could be expected to be influenced by changing travel costs.

1.4.10 Many large-scale models with a variable demand model containing any or all of the responses of trip frequency, mode choice and distribution have been built on a 12 or 24 hour weekday basis depending on time restrictions on surveys. In these cases the proportion of travel that takes place in each time period will need to be estimated separately, probably by factoring methods.

1.4.11 There are two distinctly different aspects of time of day choice; these are termed **macro time period choice** and **micro time period choice**. Macro time period choice represents the choice *between* broad modelled time periods, whereas micro time period choice represents choices *within* a modelled time period.

1.4.12 Some trip timing is, forced by changing journey conditions and times, but there may also be a deliberate choice to travel at a different time to reduce travel time and or costs. Macro time period choice occurs when travellers alter the timing of their activities and hence the time of day in which they are travelling. Micro time period choice modelling represents much smaller adjustments to departure times and packages to do this are currently under development. Peak spreading is mainly a micro response and represents travellers adjusting their travel behaviour without substantially altering their preferred arrival time or the timing of their destination activities. To date, apart from some specialist modelling, variable demand models have only included macro time period choice to represent transfer of traffic between broad time periods.

1.4.13 Macro time period choice between broader time periods would be essential if the scenarios to be examined include differential charging or other large changes in cost at different times of day. If this mechanism is included, as described in Section 1.8, then *sensitivity testing* (see *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4)) of the strength of the parameters should be used to examine the possible range of responses.

1.4.14 If modelling predicts unrealistically severe congestion in the peak hour, micro time period choice modelling to reallocate trips between the peak hour and the shoulders may be used to achieve a more realistic estimate.

1.5 Form of Models

1.5.1 An important issue that needs to be decided is the form of the demand model used for particular applications. There are a number of model forms that can be employed and these can generally be placed into three categories:

- absolute models, that use a direct estimate of the number of trips in each category;
- absolute models applied incrementally, that use absolute model estimates to apply changes to a base matrix; and
- pivot-point models, that use cost changes to estimate the changes in the number of trips from a base matrix.

1.5.2 The use of these models would depend on the compatibility of the based demand matrix if in P/A format (see Section 1.3 of TAG Unit 3.10.2) and the assignment matrix used.

- If the base demand matrix [P/A] is compatible with the base assignment matrix [O-D], in the sense that by applying time of day factors to the P/A matrix, adjusting for vehicle/person factors, and transposing the matrix to get the return trips, the required O-D matrices are generated, then (demand) changes to the base P/A matrix may be modelled either absolutely, or incrementally. In the latter case, this is done by incorporating the base values explicitly in the demand formulation;
- If the base demand matrix [P/A] is not compatible with the base assignment matrix [O-D], then (demand) changes to the base P/A matrix may be modelled either absolutely, or incrementally. In the latter case the base values are incorporated explicitly in the demand formulation, but only the implied change in demand after converting to O-D form is used to adjust the base assignment matrices.

1.5.3 The latter two methods in Section 1.5.1 retain all the detail of the observations, but generally face difficulties where too many (or key) cells in the observed matrices are empty because of the limited amount of surveying possible. This section explains the differences and the preferred approaches.

Absolute Models

1.5.4 Absolute demand models generate estimates of trip numbers, based on a model that is calibrated to fit as closely as possible to the known observed movements and the resulting model is used to directly forecast future trips. Base year and forecast trip patterns are produced independently of each other, using common model parameters. The sensitivity parameters used in absolute models should be calibrated from local data. In addition to the calibrated sensitivity parameters, however, mode-specific and movement-specific constants will usually be required to achieve an acceptable fit to the observed data.

1.5.5 The simplest approach to forecasting with absolute models is to run them separately for each combination of forecast year and scenario. Comparison of the 'without scheme' to the 'with scheme' case then gives an estimate of the impact of the scheme under consideration. This was the common practice with many traditional multi-stage models in the past. However, the fit of these models to the observed base year data was often quite poor, even where calibration constants, disaggregated by area type, had been used.

Absolute Models Applied Incrementally

1.5.6 In recent years, forecasting approaches have attempted to make use not only of the absolute model but also the 'observed' base trip matrices on which it was calibrated. This could be by factoring the forecast trip matrices by the ratio of the base year synthesised matrix to that of the base year observed matrix, so that:

Future year matrix = (Base year 'observed' matrix / Model absolute base year forecast) * Model absolute future forecast

1.5.7 However, this could lead to odd results where the cells of the observed trip matrix are zero. A way around this problem that has been used by some multi-stage models is to employ an additive approach so that:

Future year matrix = (Model absolute future forecast - Model absolute base year forecast) + Base year 'observed' matrix

1.5.8 Thus, the future forecast by the model is increased by the difference between the base year observed trip matrices and those produced in the base year by the model. The danger with this approach is that, sometimes, negative cell values can result.

1.5.9 In either of these ways the important differences between the observed matrices and the base year model that were not picked up in the calibration process are reflected in the forecasts. The first approach has been used in the recent PRISM model of the West Midlands and the second approach has been used in the Transport Model for Scotland (TMfS).

Pivot-Point Models

1.5.10 Pivot-point models estimate changes in trip patterns relative to a base matrix in which, normally, observed movements are used as much as possible. Such model applications are often described as 'incremental' or 'marginal'. The predicted relative changes are applied to the base matrix, so that the complexities of the base matrix are preserved. Where it would be difficult to calibrate a demand model to reproduce the observed pattern of travel these incremental models can be used to predict from (pivot off) this base matrix and associated costs. The matrix can also be updated in whole or in part without altering the forecasting model since the parameters controlling the mechanisms can be independent of the calibration of the base model.

1.5.11 In such pivot-point models the forecasting mechanisms estimate changes in trip patterns as a function of the observed base, rather than estimating absolute numbers of trips. With true incremental models the base year conditions (costs) and the reference trip pattern (derived from the base year trip matrix assuming no changes in travel costs) are direct inputs to the forecasting process. Such an approach has a number of benefits. It can use existing data relatively easily, and the parameters used in the model can reflect known sensitivities to changes in input variables without having to perform the additional and time-consuming task of fitting to an observed pattern. That time consuming task requires the analyst to identify differences in mode specific constants, in the case of mode-choice, and sensitivities in different parts of the network (known as calibration areas). However, by not carrying out that task, the parameters will generally need to be calibrated using external data sources, or imported from other demand models (see Section 1.11 below which discusses illustrative parameter values). Models that use this approach to forecasting are described further in *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4).

1.5.12 The recommended pivot-point approach is given below and illustrated in Figure 3, which highlights the way that the forecasts pivot off the base year costs (reference cost).

Recommended pivot-point approach:

Step 1. Growth factors are applied to the base year matrix A to produce a reference case matrix B.

Step 2. Pivot off the reference case to get to point C which is the do-minimum

Step 3. Pivot off point C to get to D which is the do-something. Note that point C represents a well converged do-minimum scenario.

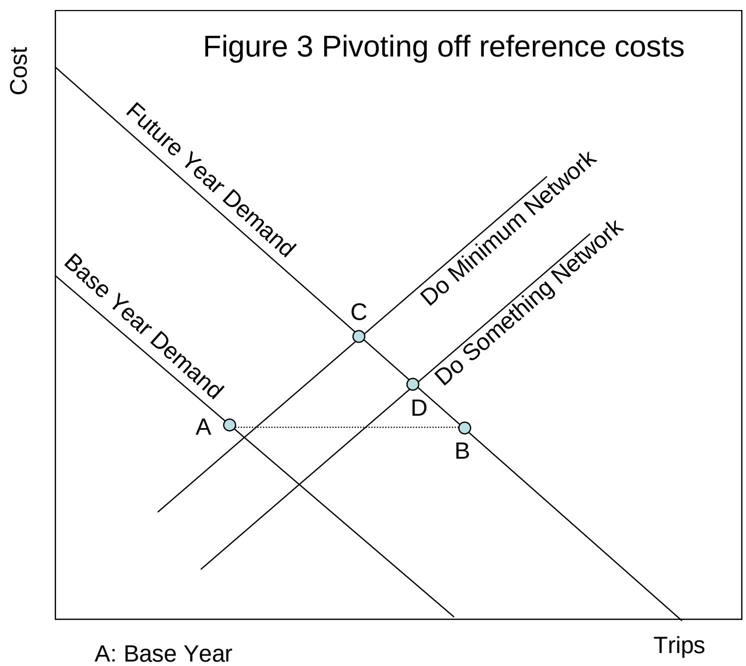


Figure 3 Pivoting off reference costs

Step 1
A->B: Furness Base year trips to future year demand

Step 2
B->C: Pivot off Reference Case to get Do Minimum or B->D for Do Something

A: Base Year
B: Reference Case
C: Do Minimum
D: Do Something

1.5.13 While this is the recommended approach to using pivot-point models, it is possible to pivot off the reference case B, to get to the 'with scheme' case D without going through C. It is also possible to use different definitions of reference trip matrix and costs, for instance, when some of the demand responses are modelled in another external model. However, these points would have to be on the future year demand line. Whatever variation of this approach is used details should be given in the forecasting report.

1.5.14 There is a number of proprietary software packages available that will undertake the required VDM forecasts but the user has to choose the underlying costs matrices and the reference trip matrix to enter into the software (see *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4).

Choice of Model Form

1.5.15 In summary there are a number of different approaches to producing forecasts of future travel demand. On the one hand there is the pivot-point method that uses trip matrices directly and forecasts by estimating the impact of incremental changes in the travel costs between a base situation and the forecast year/scenario. In most cases the base year costs are used as the pivot-point for the forecasts but in some circumstances other pivot-points may be used. On the other hand there are absolute models where the forecasts make no direct use of base year trips or costs. The approach of using the observed trip matrix to adjust absolute forecasts (Incremental Forecasting Approach) has blurred the distinction between forecasts produced from absolute models and incremental models. Whilst all pivot-point approaches are incremental, absolute models may be used directly or with an incremental adjustment.

1.5.16 Whilst the usefulness of the pivot-point methods of being able to make use of any peculiarities of the observed trip matrix is important, an equally important issue is the relative reliability of any observed trip pattern compared with that of an absolute model estimate. Both are limited by the expense associated with attempting to observe all the cells of the matrix, covering all the dimensions (of mode, time of day, on the one hand, and purpose and segments on the other). In addition, there are questions of temporal variability (seasonal, day-to-day, etc.). Thus, except in models that have very large zones, 'observed' matrices drawn from survey data are likely to contain a very high proportion of cells with zero trips. Without importing synthetic data an incremental model will never fill these cells. If the travel that the cell represents is truly infeasible then zero trips may be a reasonable assumption, but if the cell is empty merely because of sample limitations some matrix infilling may be advisable. These infilling techniques are discussed in detail in DMRB Volume 12.1.1.

1.5.17 In contrast, traditional absolute forecasting models do not use the local observed data directly but use it only for calibration purposes. Advocates of their use argue that this provides greater confidence that the model is properly fitted (in behavioural terms) to local conditions, and that estimates can be made of the statistical accuracy of the model. Additionally, estimates are automatically produced for all movements including those that pass through the survey points, but were zero because of sampling variation.

1.5.18 However, in practice, often a large number of calibration factors need to be used with most absolute models to provide a reasonable fit to the observed data, often in a rather arbitrary way especially in the case of distribution models. In addition the parameter values so obtained may have arisen from some masking of observed differences (which may for example be distance-related effects) by the zonal or mode specific constants and may therefore either underestimate or exaggerate the true sensitivity of travellers to changes. Consequently, the apparent superiority of local calibration is often not fulfilled in practice except in large-scale transport studies, where the data collection and calibration can be sufficiently comprehensive.

1.5.19 Pivot-point approaches are attractive, but it is necessary to sound a note of caution. In the first place, any deliberate decision not to attempt to synthesise observed cross-sectional variation has potential forecasting implications. To the extent that a model is deficient in synthesising, it may be equally deficient when used incrementally. This is called model mis-specification.

1.5.20 However, the observed data are themselves affected by sampling error considerations. One way around this problem is to use a mixture of observed and synthesised data, even for movements supposedly wholly observed, particularly where sample sizes are small and sampling error consequently high. The relative quality of the data and the synthetic model need to be considered, especially if long-term redistribution processes might not have been fully realised. Usually the potential model mis-specification errors of distribution models are such that the observed data should be given greater weight in any mixed base.

1.5.21 However, by taking a weighted average of the observed and synthetic matrices empty cells are eliminated and greater weight is given to cells where there are more observed trips than expected from the locally-calibrated synthetic model. Relative weights should reflect the relative accuracy of the two forms of estimates. If these are not known then a rough guide would be to use 90% of the observed estimate and 10% of the synthesised estimate. Synthesised data is also required for movements not captured by the survey strategy.

1.5.22 The main problem with using pivot-point models occurs when the base matrix contains few or no trips for a set of movements, but the forecasts expect large changes in these movements to occur. This often arises when a zone is re-developed or has no trips to or from it in the base situation. In these situations the forecast will have to be synthesised exogenously for these movements.

1.5.23 In practice, the base matrix may have elements of both observed and synthetic approaches as some parts of the 'observed pattern' of travel may have to be synthesised from other observed data. This is most obviously true when the partial matrix technique is used to fill unobserved cells in the observed OD matrix (an absolute demand model may be used to estimate the likely values for empty cells), but the synthesising of various combinations of household structure from incomplete information, especially for forecasts, is another example.

1.5.24 The Department’s preference for road-scheme appraisal is to use an incremental form of model whether pivot-point or based on incremental application of absolute estimates, unless there are strong reasons for not doing so. Such reasons could include situations where there are large changes in land use between the base and forecast years, which will significantly change the distributions of origins and destinations.

1.6 Which Modes at What Detail?

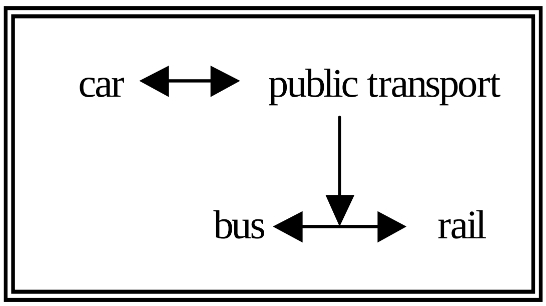
1.6.1 This section discusses the representation of different modes in the variable demand model. Usually, the main alternative mode to car will be public transport in its various forms. In some circumstances it may be desirable to represent competition from walk and cycle; however, in most highway assessments explicit modelling of slow modes is unlikely to be worthwhile, especially in inter-urban areas. It may be acceptable to include the alternative mode(s) merely as a set of fixed costs. Where mode-choice is important to the assessment, it may be necessary to model the journey components in detail, including the effect of changing road conditions on bus travel. Further advice on public transport modelling will be released as a series of TAG Units in due course.

1.6.2 As Section 1.4 explained above, it is almost always desirable to include some representation of modal choice in variable demand modelling, but the level of detail depends upon the importance attached to it.

Public Transport

1.6.3 It is likely that the competitive public transport mode will be rail, or intercity coach, for longer distances, and bus or LRT for shorter urban distances. In some cases more than one mode may offer a competitive alternative to car, in which case the demand model should include a higher level car/public transport modal split mechanism, with a separate split between the available public transport modes below this in the hierarchy (this is known as a nested approach - see Section 1.9). That nested approach avoids the so-called “red bus/blue bus problem”, in which merely splitting one existing mode into two identical ones will apparently predict mode shift when using a multi-nomial logit model. Nested models should have at least two levels, with choice between private and public transport at the higher level, and then between different public transport modes at a lower level. The choice sensitivity lambda parameters should be larger at these lower levels than at the higher level. If there is a separate public transport assignment model, then this secondary choice can be made in the assignment stage, with (in that case or when using a nested approach) the generalised costs fed back to the higher-level mode choice as flow-weighted or composite averages across the public transport modes. Detailed advice on these issues will be released as part of the public transport modelling guidance to be released in due course.

Figure 4 Example of a nested choice structure



1.6.4 If there is little real competition between public transport and car, the public transport generalised cost estimates can be made with limited precision. The level of competition can best be judged from local knowledge of modal split for car-available travellers. This can only be obtained by travel surveys, which set out to identify car availability: in most cases it will have to be proxied by household car ownership, but this will generally overestimate true availability. Approximate values for an area can be obtained from the 2001 Census.

1.6.5 As a general guide, if public transport is chosen by less than 5% of travellers, use of fixed public transport costs will suffice, unless public transport alternatives need to be assessed as part of the scheme appraisal.

1.6.6 Consideration also needs to be given to the future role of public transport. If public transport is expected to play a significant role in the future (either through specific scheme or policy implementation) then, irrespective of current mode choice, development of a mode choice model may be appropriate.

1.6.7 Where more detailed assessment of the impact on public transport will be needed or it is known that changes to the public transport network will be assessed, the reader is referred to the guidance on Public Transport Modelling and Forecasting for further advice.

1.6.8 Rail travel can be represented by a fixed-cost matrix even if it takes a larger than average share of car-available travel, since it is not subject to road congestion. However, an equivalent effect is rail overcrowding as demand changes, but the effect is difficult to estimate and is often ignored. Access to the stations is likely to be by car for some trips, and a mean generalised cost of access across all relevant modes (walk, bus, and car including parking charges) should be estimated. Large changes in demand for either rail or bus might result in changes in service frequency, but these effects are best examined using a specialised public transport model.

1.6.9 If the modal competition is considered important, then a more detailed representation and costing of the public transport alternative will be required. The reader is referred to the guidance on Public Transport Modelling and Forecasting for further advice.

1.6.10 Where public transport is an important component of the transport scheme, and a detailed public transport network (assignment) model is needed the generalised costs for each OD can be extracted directly from it. However, the access costs to each mode may need to be added. Competition between different public transport services, or modes, can also be modelled in assignment. Where an existing highway assignment model is to be used in conjunction with an existing public transport assignment model, the generalised costs of the two models need to be consistent with one another.

1.6.11 If the scheme being assessed affects the journey times of the public transport services appreciably, this should be estimated. This will usually only be the case for bus services, which may be speeded up or hindered by the scheme, and which may also be affected by changes in traffic congestion. In practice, there is a low propensity of car users to see bus as a realistic alternative, except perhaps for park and ride schemes.

1.6.12 A worked example of estimating modal split and the effect of changes in road speeds on bus users is provided at Appendix 3.

Slow Modes

1.6.13 Mostly, assessment of a road scheme is unlikely to be affected by whether slow modes are included or excluded from the modelling. Walking from car parks can be modelled in the car assignment process. Walking for the whole journey only offers appreciable competition to the car for short journeys, typically no more than a kilometre or so, and while cycling can be competitive over much longer distances (and especially so where traffic congestion slows the car), it is rarely considered an acceptable choice by the majority of car-available travellers. This is not to argue against the merits of the slow modes which are environmentally friendly and healthy, but merely to treat them realistically in transport demand modelling, especially when trips are too long to be feasible by slow modes.

1.6.14 There may be special reasons for examining the role of slow modes in the modelling (for instance for schemes on radial routes in urban areas with high cycle usage) and their inclusion may well affect predictions of short trips on a road scheme. If they are included, they will stabilise total trip rates, and there is less reason to model trip frequency. This is because when the costs of mechanised travel fall, as speeds increase or operating costs fall, there will be modal transfer from slow modes to mechanised modes. This appears as the trip frequency response in a model lacking the slow modes, but as modal shift in a model which contains them. Over time, the inverse effect may be as important. When the costs of congestion increase, or operating costs rise (perhaps due to congestion charging or increases in parking charges), there may be modal transfer from car to slow modes. This appears as “trip suppression” in a model lacking the slow modes, but as modal shift in a model which contains them.

1.6.15 When slow modes are omitted trip frequency elasticities should be stronger, since then it has to represent the effect of this “slow modal” transfer. If slow modes are to be included, then a cost-responsive trip frequency mechanism can be omitted altogether. If they are treated as a separate mode, as opposed to them being included in a general non-car mode, it will normally be adequate to treat their generalised costs as linearly dependent on OD distance travelled, via an average speed that is conventionally 4 km/h for walking and 12 km/h for cycling. Walking speeds in particular are a function of the number of roads crossed, and the amount of traffic on these roads, but this aspect is rarely captured in modelling.

1.6.16 Slow modes may be included in mode split at either level of the hierarchy, as part of a higher level car/public transport/slow split with perhaps a sub-modal split between walk and cycle, or, more usually, the higher level split may be kept binary between car and public transport plus slow, and the latter split at the sub-modal level.

1.7 Trip Distribution

1.7.1 Trip distribution models spread the forecasts of generated trips over the available destinations, depending on the generalised cost of reaching that destination. Most distribution models are designed to guarantee that the total number of trips from the origin zone (or to the destination zone) is equal to the total number of trips for that zone forecast at the trip generation/frequency stage. If the design guarantees that property for both origins and destinations the model is known as doubly constrained.

1.7.2 Distribution models can be applied in terms of zonal productions and attractions or less satisfactorily in terms of origin and destination trip totals. *VDM Scope of the Model* (TAG Unit 3.10.2) discusses the distinction and the implications for the modelling process. Various functional forms are available, but most models in this country have used a logit-type formulation to represent the influence of travel costs on choice of destination, in a similar way to choice of mode etc. It is common to use doubly-constrained models for forecasting commuting and education trips, so that each zone attracts and generates a fixed total of work trip ends, and singly-constrained models for other purposes, where only the total number of trips originating in each zone is fixed, using the techniques described below.

1.7.3 In addition to cost, distribution also depends on some measure of the attraction of a zone, estimated in terms of the numbers of “opportunities” such as jobs or retail floorspace in the zone. These reflect the likelihood that the zone will be chosen as a destination, other things being equal (though in doubly-constrained models the attraction is simply the number of trips required to end in the zone).

1.7.4 Whilst trip distribution models were originally expressed as “gravity models” (using a negative power of distance) they are now usually formulated as a logit model, as underpinned by random utility theory for discrete choice modelling and because of its mathematical tractability. Other functions such as simple power, Tanner or Box-Cox functions are available as well as the commonly used exponential formulations (see *VDM Appendices* (TAG Unit 3.10.5)).

1.7.5 The main stumbling block in their use lies in estimating the trip attraction factors for each zone in a robust and reliable way, and in determining parameters which have real predictive values. This is difficult, since the distribution parameters are normally calibrated to recreate the (cross-sectional) data observed at a given point in time, which depend on a wide range of historic and socio-economic factors, which cannot be captured fully in the modelled transport factors. Those historic factors can be large enough sometimes to mask the true choice process amongst closely competing destinations in an equilibrium model.

1.7.6 Consequently, the models’ ability to predict choices and changes in trip patterns due to changing transport factors is generally unproven. For this reason **it is recommended that trip distribution models normally have an incremental form, building on a largely observed base. Local parameter values should be calibrated for use in the model; however, if there is insufficient observed data for satisfactorily calibration, externally derived parameter values should be used**, although some adjustment may be needed to deal with any under-representation of competing destinations and situations where major changes to land-use are expected to occur.

1.7.7 Predictions of trip distribution are usually “Production Constrained” to a total based on forecasts of trip-ends (see *VDM Scope of the Model* (TAG Unit 3.10.2)). Similarly, the trip matrix can be constrained to match a required number of total trip attractions. In general, the trip matrix and productions will be disaggregated by trip/traveller segments, and will have to satisfy the constraints within each individual segment, such as each trip purpose or traveller type.

1.7.8 There are four main decisions that have to be made about the use of trip distribution models within a variable demand model:

Production/Attraction or Origin/Destination Modelling

1.7.9 The implications for this are discussed in *VDM Scope of the Model* (TAG Unit 3.10.2) and the choice is likely to be made based on the availability of data, what other demand responses are being modelled, and what form of demand/assignment model is being used. However, when building a new model, or substantially updating one, the presumption should be that any new matrices are assembled and used as P/A defined trip matrices.

Doubly or Singly Constrained

1.7.10 In general, doubly constrained models should be used for commuting and education. This reflects the relative confidence in the measures of attraction (employment and student numbers) for commuting and education trips, as well as the relatively fixed nature of these attraction values in the short term.

1.7.11 Other purposes such as shopping, social and leisure trips are typically modelled as singly production-end constrained. For these purposes, the trip end factors reflect the attraction of destinations, not the actual numbers of trips attracted and ideally the availability of intervening similar destinations between the origin zone and the zone in question. In practice the required estimates need be only relative, and usually depend on a weighted combination of quantities like shopping floorspace or employment, with the weights obtained from fitting regression models, or they may be obtained from trip-end models such as TEMPRO.

Incremental or Absolute (based on wholly synthetic models) Forecasting Models

1.7.12 Where possible incremental models should be used, since these usually have the benefits of a more directly observed trip matrix (see Section 1.5 for a discussion of this issue). This matrix reflects not only the pattern of trip ends and the costs of travel between them, but also the cumulative impact of past travel and settlement decisions - which an absolute trip distribution model, based on current costs and trip patterns, cannot take into account except via the inclusion of a multitude of area-specific fitting constants. These special factors (usually known as 'K' factors) represent that part of the interaction between zones that does not conform to the general synthetic model expectations. In calculating those factors it is advisable to first identify calibration areas and then vary the distribution parameter by calibration area as well as traveller type before resorting to such zone-to-zone factors. None of this is necessary when the model is incremental.

Model Form

1.7.13 There are a number of different model forms suitable whether the model is doubly or singly constrained, incremental or absolute. It is expected that the model form will generally be logit; however, a number of different deterrence functions are possible. These are discussed more fully in Appendix 2. Starting with an absolute model formulation since they are easier to understand (discussion on incremental models can be found in appendix 4), the trip distribution model can be written in a number of equivalent forms such as the general form:

$$T_{ij} = a_i \, b_j \, O_i \, D_j \; f(G_{ij})$$

Here a_i and b_j are simple or iteratively calculated “balancing factors”, whose values depend on trip end constraints. In multi-stage models it is clearer if this general form is expressed as a destination choice model.

If the distribution model is singly (origin)-constrained the equivalent destination choice model is:

$$T_{ij} = \frac{O_i D_j f(G_{ij})}{\sum_k D_k f(G_{ik})}$$

which satisfies the origin constraint:

$$\sum_j T_{ij} = O_i$$

If the distribution model is doubly-constrained, the destination choice model is:

$$T_{ij} = \frac{O_i b_j D_j f(G_{ij})}{\sum_k b_k D_k f(G_{ik})}$$

where the b_j's are calculated iteratively to satisfy the destination constraint:

$$\sum_i T_{ij} = D_j$$

In the case of origin constrained trip distribution D_j is some function of the attractiveness of destination zone j, and in the case of a doubly-constrained trip distribution model O_i and D_j represent total origin and total destination trip ends respectively. G_{ij} represents the generalised costs of travel between i and j and $f(G_{ij})$ the deterrence function which may or may not contain (multiplicative or additive) K_{ij} factors.

1.7.14 In the above equation there are a number of different deterrence function forms that can be adopted for $f(G_{ij})$. In a true gravity model the deterrence functions are power functions $f(G_{ij}) = G_{ij}^{a}$ (and originally interzonal distance was used instead of G), but it is standard now to use an exponential form:

$$f(G_{ij}) = \exp(-\lambda_{dist} G_{ij}),$$

or with multiplicative K_{ij} factors:

$$f(G_{ij}) = K_{ij}\exp(-\lambda_{dist} G_{ij})$$

or with additive K_{ij} factors:

$$f(G_{ij}) = \exp(-\lambda_{dist}(G_{ij}+K_{ij}))$$

1.7.15 The calculation of costs should use composite cost G_{ijcomp} (see Section 1.9) calculated only across the stages lower in the hierarchy of mechanisms.

1.7.16 Different values of the distribution parameters can be used for different cells, or over different cost bounds, or a completely empirical relationship between deterrence and cost can be used. However, since most of the evidence on suitable parameter values relates to the logit form, and this form is the most common for other demand responses, this should be the first choice. Where alternative parameters are justified by a study of the base situation, the logit parameter value may vary by origin or destination zone. However, any logit or exponential distribution model implies that the sensitivity to a given absolute change has the same effect on travel between zones far apart as on those close together, and this sometimes leads to large percentage changes in long-distance trips. This can be mitigated by careful choice of calibration areas. It should not be a problem for local models, but could give rise to unusual forecast changes for models with very long and very short trips (though trip-end constraints will mitigate the effect somewhat for doubly-constrained models). Where locally derived parameters have been produced by calibration area, then the trip matrix may need to be split into categories based on these calibration areas (for instance trips to an urban centre) before forecasting is undertaken.

1.7.17 There will not normally be a requirement to model trip frequency for doubly-constrained trips, since the constraints on total travel are usually assumed to be binding. This normally implies a constant frequency of travel to work, with changes limited to the choice of mode and destination. This implication does not hold if slow modes have been omitted and they may form a sizeable percentage of commuting trips. In these circumstances a trip frequency response could be added although care would need to be taken that the response is not simply affecting long distance trips rather than the short distance trips the response is acting as a surrogate for.

1.7.18 Shopping, social and leisure trips are typically modelled as “Singly Production Constrained”. For them, the trip end factors reflect the attraction of destinations, and not the actual numbers of trips attracted. This is not especially easy to estimate, since it will depend not simply on the amount of the activity available in the zone (for example, the retail floorspace) but also on its type and quality, and on the availability of intervening similar destinations between the origin zone and the zone in question. For some of these purposes it may be logical to consider a trip frequency effect on top of the distribution effect; that is decreasing costs will lead to greater numbers of trips of that purpose as well as change the destinations. An example of this is leisure or holiday trips, but shopping trips too are likely to be elastic, especially if the model does not include slow modes, since walk trips to the local shop may become mechanised trips to more distant shopping centres if travel costs fall.

Spatial issues

1.7.19 Trip distribution models are likely to be the demand responses most sensitive to the spatial extent of the model area but the degree of sensitivity will also depend on the form of the distribution model chosen. Three issues are worth highlighting with respect to trip distribution models.

- Where possible, all likely destinations for zones within the main area of interest should be modelled. This is particularly important for trip distribution models since trip increases in one area, say within a corridor of interest after improvements, should lead to decreases to other destinations. This will have implications for traffic quantities and benefits (overall, and within given areas).
- Average intra-zonal trip costs should be calculated as accurately as possible to remove bias against shorter trips in the distribution model. The modelling of intra-zonal trips is especially important and usually needs to be considered separately, as most assignment models do not assign intra-zonal trips and hence no costs are output. Power function models are particularly sensitive to very low intra-zonal costs, and where mode-choice is undertaken lower down the hierarchy than distribution the distribution of car trips using a power function could lead to an excess of very short-distance car trips.
- If destinations outside the main study area are potential alternative destinations then the costs to these destinations should be calculated reasonably accurately, even if the network, and the zoning system outside the study area, is of a coarse nature.

1.8 Time of Day Choice

Macro Time Period Choice

1.8.1 Macro time period choice, involving the transfer of trips between broad time periods, can be modelled as a logit choice in a similar way to the choice mechanisms described for the other stages of demand modelling. However, if the demand modelling uses the typical division of time into two peak periods and an inter-peak (see *VDM Scope of the Model* (TAG Unit 3.10.2)), the freedom of most trips to transfer between them will be severely constrained: few work trips, for example, could move outside the three-hour peak periods entirely, and such a mechanism might be applied predominantly for shopping as opposed to the journey to work.

1.8.2 To model macro choices, it is necessary to know what proportion of each type of trip takes place in each period. At a macro level trips must be allocated to a discrete time period even those which start and finish in different periods. An incremental logit model can then be used to modify the total number of trips **of each type of trip** in each time period according to the changes in the mean generalised costs in each period.

1.8.3 **Macro time period choice should be considered when strong cost differentials between time periods are expected to develop or change.** This is obviously the case where different charges are introduced for use of a road, rail or bus service in the peak and inter peak or off-peak, or where different levels of access to road capacity are being contemplated, or perhaps where peak surcharges are introduced for parking in a way which affects a large proportion of traffic. In these cases it is obviously important to choose the modelled time periods to facilitate the modelling of the differential costs.

1.8.4 There is limited evidence on the strength of the macro time choice mechanism. Recent Departmental research,(http://www.dft.gov.uk/stellent/groups/dft_econappr/documents/divisionhomepage/040158.hcsp) suggests that time period choice is generally more sensitive to changes in travel conditions than mode choice. The report also concluded that the choice of time period is generally more sensitive for changes between short than long periods: however this finding cannot be extended to 15 minute periods on the basis of the tests made in the study...

1.8.5 Where circumstances justify the use of this technique, it can be important to apply different sensitivity parameters to different trip purposes. For further details see analysis in (http://www.dft.gov.uk/stellent/groups/dft_econappr/documents/divisionhomepage/040158.hcsp)

Peak Spreading (Micro Time Period Choice)

1.8.6 It is common experience that when demand grows in a congested network the peak in demand tends to occupy a longer time. The peak is unable to grow higher for lack of capacity, so additional demand is accommodated in the shoulders of the peak. This effect is known as "peak spreading", but it occurs because of a mixture of responses, both involuntary and voluntary:

The length of the peaks will spread as congestion grows, because lower speeds mean that any given journey will take longer to complete and will occupy a longer period. The traveller has little influence in this, and the delays caused by this effect are often represented within the assignment modelling itself, as described in *VDM Scope of the Model* (TAG Unit 3.10.2).

Travellers can deliberately change their time of travel, departing and arriving earlier, or later, than their preferred time. It is common experience in congested conditions that a quarter hour change in departure time can change the expected mean travel time significantly. Some travellers will find such a change acceptable, because the saving in journey time outweighs the benefit they attach to arriving at a preferred time. The response is clearly more available for travellers who have some flexibility in precisely when they must arrive and is applicable to those work trips that have some degree of flexibility for earlier or later arrival.

1.8.7 In the face of increasing congestion, some travellers will adjust their departure times or arrival times to gain a reduction in travel time. In principle, this can be represented as a choice mechanism reflecting the generalised cost of travel to which has been added the cost of not arriving at the preferred time: this is a "schedule disutility term", essentially an extra component to the generalised cost which measures how far the actual arrival time is before or after the preferred time.

1.8.8 When a micro time period choice response (that will include a schedule disutility term) is included in a variable demand model, evidence suggests that this is likely to be the more sensitive than other responses except route choice.

1.8.9 DfT has developed a departure time choice model (HADES) which interfaces with standard assignment packages (Van Vuren, 2002). This represents a continuous range of departure times and interfaces with a range of assignment software which use a small number of time periods. For further information see www.dft.gov.uk/stellent/groups/dft_econappr/documents/divisionhomepage/032182.hcsp. The HADES model is being developed within the DIADEM software (see *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4) for further details) and is expected to be available in due course. In the meantime further advice on the use of HADES in particular local applications should be sought from the DfT.

1.8.10 Other approaches can be used to represent peak spreading between the peak hour and shoulders, such as multinomial logit, although there are theorectical reasons why this form of model may not be a reliable predictor of choice between shorter time periods. Some of these techniques are described in DMRB 12.2.1. These may be suitable in particular local applications until the DIADEM package offers an appropriate micro time period choice facility.

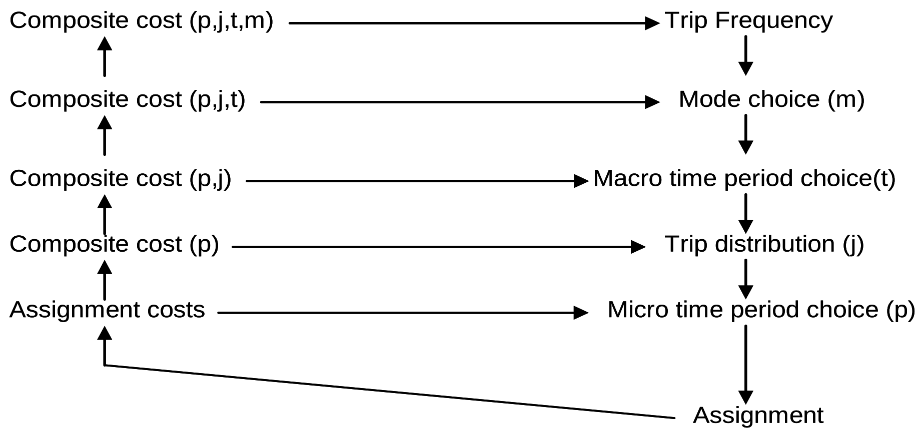
1.9 Hierarchy of Responses and Composite Costs

1.9.1 Once decisions have been made on which responses to include in the model, the hierarchy in which those responses are considered must follow certain rules. This is not simply a question of mathematical or computational convenience. The sequence of the mechanisms is important to the overall outcome, and the resultant elasticities of demand and the predicted travel pattern will be affected by it. This section describes how the hierarchy should be determined.

1.9.2 The appropriate hierarchy or sequence of choice mechanisms must be determined by the relative sensitivities (the lambdas of a logit model) of the choices to the generalised costs or disutilities of travel. Different sequences for different purposes and/or segments of the travel market are often appropriate.

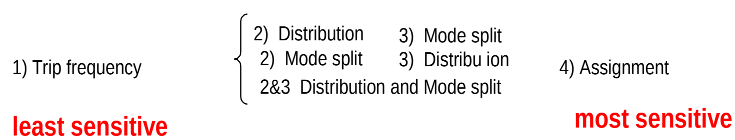
1.9.3 A mechanism placed "higher" in the hierarchy of demand mechanisms should reflect the "composite cost" of choices lower in the hierarchy and allow for how a choice with high costs is unlikely to be chosen. A "logsum" of costs as described later in this section has that property, but requires "higher" demand mechanisms to have a smaller sensitivity, to avoid a plausible change in generalised cost producing an implausible shift in demand.

1.9.4 The sequence of calculations is that, during each cycle, the composite costs must be calculated for each level in the hierarchy, since each level refers to different combinations of choice lower in the hierarchy.



1.9.5 Thus the cost calculation starts at the bottom of the hierarchy and works its way up the levels, adding one more choice into the composite cost at each level. The choice calculations are then made down the hierarchy and the whole cycle is recalculated in the next iteration until an acceptable degree of convergence is achieved. An example sequence of mechanisms is shown below; however it is unlikely that a model would be developed in this form and the mechanisms may need to be positioned at different places in the hierarchy, depending on their relative sensitivities.

1.9.6 Available evidence suggests that the sensitivity of trip frequency is very much smaller than for the other mechanisms, and it is justifiable to always treat this choice as first in the hierarchy. That argument and the discussion set out below leads to the following practical hierarchies. However, if the model represents slow modes, there is less need to make the trip frequency stage responsive to *local* travel costs. The hierarchy is therefore likely to be one of the following (excluding time of day choice responses):



1.9.7 ***Route choice*** is invariably modelled as the most sensitive response below the other demand mechanisms. In equilibrium, there is little or no difference in utility or generalised cost between the routes which are likely to be used for any given origin-to-destination journey; if costs change, a new equilibrium involving some change of route between the minimum cost alternatives is quickly established. Thus the route assignment part of the modelling can be considered separately, though of course demand and assignment must be treated iteratively to obtain an equilibrium solution (i.e. the Wardrop equilibrium).

1.9.8 Where ***macro time period choice*** is thought applicable (where there is expected to be differential changes in the costs or capacity in different time-periods) evidence suggests that the position of this mechanism is at a similar level to main mode choice for most purposes.

1.9.9 Where ***micro time period choice*** is being modelled, this should be placed above assignment but otherwise at the lowest level (most sensitive) of the hierarchy.

1.9.10 The main decision centres on the relative positions of distribution (destination choice) and mode split (choice).

1.9.11 The distribution model should precede the mode split if the distributional parameter (lambda) is smaller than that for mode split, and mode split preceding distribution if the opposite is true. If the two are similar, within the uncertainties that are likely to be relatively large, there is a case for simultaneous calculation using a single sensitivity (lambda) value for each traveller segment.

1.9.12 However, if for example, destination choice has a larger sensitivity parameter than mode choice, yet mode split was mistakenly calculated after distribution, an increase in the cost of, say, car travel might increase the mean (composite) cost of travel on which distribution is based. In extreme cases that could shorten all trips to such an extent that not only is car use decreased, as required, but also travel on the competing modes, which is implausible. (Such an effect is often described as a perverse cross-elasticity.)

1.9.13 In the multinomial logit formulation described in Appendix 2, a given level of the main choice hierarchy may split the choice into separate sub-levels. Such "nesting" is most often met in mode choice, where the split between competing public transport modes is made at a lower level than the primary split between car and public transport (and possibly the slow modes also). This avoids the "red bus/blue bus problem" where separating the bus mode into two without nesting apparently affects the predicted total bus share. Nested logit can be applied to other demand mechanisms; however, as for example in a time-period split between broad peak and off-peak periods, and then subsequently between narrower periods within the peak. Nesting can also use high-level large zones in distribution, and subsequently a further distribution within the larger zone to finer zones, but this is likely to be relevant only to more specialised models than the ones addressed by this advice.

1.9.14 Where sufficient local data of suitable quality exist, and the skilled resources required are available, lambda values should be estimated (calibrated) from that data and the hierarchy selected so that the less sensitive of the two responses is positioned above the more sensitive.

1.9.15 Where suitable data or resources are not available to permit parameter values to be estimated, then it may be possible to select the hierarchy on the basis of local knowledge about the relative sensitivities of destination and modal choice (from existing local models, for example, where the lambda values have been estimated and the adopted choice hierarchy has been justified).

1.9.16 If there is insufficient local data or resources for estimation, and no suitable local model from which the parameters can be transferred, it will be necessary to consider the illustrative values provided from Section 1.11 as the basis for the choice hierarchy. All the models used to derive the parameter values quoted in Section 1.11 were rigorously calibrated against local data and all showed that main mode choice was less sensitive than destination choice. In the absence of any information to the contrary, this is therefore the hierarchy which should be adopted.

1.9.17 Whichever approach is adopted, it is essential to apply "realism testing" to a broad range of transport changes (see *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4)) to ensure that the model responds rationally and with acceptable elasticities.

Composite Costs

1.9.18 Unless mechanisms at two adjacent levels in the hierarchy are calculated simultaneously (which is likely to apply only to the mode and destination choice stages), it is necessary to formulate a composite cost or utility across the most sensitive (or lower) choice to use as an "average" in the least sensitive (or higher) choice calculation. This cannot be an arithmetic average, since it is clear that where a choice has high costs and is unlikely to be chosen it should be given little weight in the composite cost. Various forms of composite cost have been used in the past – see for example Senior and Williams (1977) – but the following, known as a logsum, is the appropriate formulation where logit models are used to determine the choices in an absolute choice model. The general formulation of the composite cost to reflect the costs faced by travellers given their previous choices lower in the hierarchy is as follows:

$$G_{comp}^{y-1} = -\frac{1}{\lambda_y}\ln\left(\sum_x \exp(-\lambda_y G_x^y)\right)$$

Where:

$G_{comp}^{\ y-1}$ is the composite cost or disutility summed over the choices x in stage y

G_x^y is the disutility or generalised cost of choice x given choice y

(for example, the stage y may refer to 'destination choice', while x varies over the destination zones

λ_y is the choice sensitivity parameter for choice stage y.

1.9.19 For example, if mode split is less sensitive than (above) distribution in the hierarchy, then the composite cost of car travel from zone i is obtained from the logsum of travel by car (choice y) to all the possible destination zones (choice x). There will be fewer trips to destinations with high travel costs, but the exponential weighting means that they will contribute little to the total composite cost. At the lowest level, the absolute composite costs should also be weighted by the zonal attractivities. If distribution is less sensitive than (above) mode split, then the composite costs used for distribution will be the logsum costs across the available modes for each origin-destination pair.

1.9.20 As trip frequency is invariably the least sensitive response, in that case for each origin the summation in the composite cost must be across all destinations, modes and time periods if those choices that are being represented in the model. If time-period choice is included, then the composite costs should include trip-weighted sums across the time periods.

1.9.21 Where the model is incremental, the mathematical form of the logit function requires that the logsum be weighted by the choice shares in the logarithmic summation. The formulation to be used is then:

$$\Delta G_{comp}^{y-1} = -\frac{1}{\lambda_y}\ln\left(\sum_x \frac{T_x^y}{T_{tot}^y}\exp(-\lambda_y \Delta G_x^y)\right)$$

Where:

T_x^y is the number of trips choosing x at stage y

T_{tot}^y is the total number of trips available at stage y.

1.9.22 If the same lambda value applies to both distribution and mode choice calculations, then both sequences of calculation, distribution-modal split or modal split-distribution, are mathematically equivalent to simultaneous calculation, where the logit split would be across all possible combinations of destination and mode, and it is not necessary to calculate composite costs from one choice set when considering the other choice set. However, if there were other responses above these two in the hierarchy, say trip frequency or time-of-day choice then the logsum of the mode-choice and distribution choices will still need to be calculated to give the correct ‘costs’ for these higher level responses.

1.10 Local Calibration of Demand Models

1.10.1 As explained in paragraphs 1.9.14 to 1.9.16, there are three alternative approaches to choosing the parameter values that control the travel responses:

- use local data to calibrate parameter values;
- use parameter values obtained from other local models;
- use "illustrative" parameter values based on general modelling experience.

1.10.2 This section provides advice on the first of these alternatives. If, after considering the issues below, it is impractical to calibrate local values, then consideration should be given either to importing values from existing locally calibrated models or to using the illustrative values given Section 1.11.

Using Local Data to Calibrate Parameter Values

1.10.3 Calibration of the parameters in the demand response mechanisms can be a very time-consuming and expensive phase, and for smaller schemes the alternatives of using other local model parameters and/or standard illustrative values should be considered.

1.10.4 Calibration of the different demand responses varies in both the amount of data required and the ease of the calibration itself. In some cases the surveys used for calibration of the model can be used for other purposes (such as eliciting behaviour in response to tolling or parking restraint), so making the necessary survey work more cost-efficient. Demand model calibration complements the equivalent estimation in the assignment phase of the correct ratio between time and money costs in the assignment of traffic to the network.

Mode Choice

1.10.5 The practicality of local calibration of the mode choice mechanism depends on the quality of the data available and the ability to distinguish between public transport travellers with and without a car available, since these two categories of traveller will have very different choice-sets. It will also depend on the extent to which choice of mode is exercised by car available travellers. These data, coupled with estimates of times and costs by the various modes using standard values of time, can enable mode-choice parameters to be estimated.

1.10.6 The quality of data about trip purpose will determine how disaggregate a model can be estimated. That is likely to be most problematic for public transport data. Whichever approach is used, a check should still be made that the model reproduces the modal-split correctly for the important movements. Similar considerations apply when using parameters from other local studies such as a regional multi-modal study.

1.10.7 Relatively simple mode-choice calibration can be undertaken where the number of car trips and public transport trips for which a car is available are known for a large number of important flows, together with estimates of times and costs by the various modes and appropriate values of time. This has been a common approach by the multi-modal studies where detailed transport modelling, as opposed to strategic models, has been required. It is important to distinguish between travellers with and without a car available for their journey. This data can only be obtained by personal survey, either in the household or on public transport, though even then the travellers' claim to have a car available often ignores competition for the car within the household. As always, the quality of data on trip purpose will determine the disaggregation possible.

1.10.8 An alternative mechanism would be to use Stated Preference (SP) surveys to estimate the important determinants of mode-choice. SP survey work provides a useful approach when considering the introduction of states that are not present in the current situation, such as tolling or parking restraints, or where new modes (to the area) are being considered. In these cases the surveys are more likely to be geared towards estimating the relative impact of items that make up the definition of generalised cost than to provide evidence of the parameters controlling the mode-choice. Specialist advice should be used to establish the mode specific constants of the relative attractiveness of the new modes. These issues will be dealt with in more detail by public transport modelling advice soon to be released as a series of TAG units. Where household surveys are being undertaken to collect other data for modelling purposes, it may be possible to attach an SP study at marginal cost. The software available to calibrate models using SP (and other data) can handle a variety of forms of hierarchy of travel responses. The output statistics can help to shed light on the most likely choice structure. This is particularly important for mode-choice, where nested choice structures are often required but this will be a rare requirement for the bulk of road schemes. Thus, one can estimate models where a traveller chooses from all available modes at one level or chooses between say car travel and public transport and then makes a subsequent choice between bus and rail or any other such mode if the circumstances require such a detailed modelling of the demand for public transport.

1.10.9 One of the characteristics of calibration using SP methods is that the results tend only to give the relative importance of different modes and their attributes, and may not reproduce current market shares, without using a "scaling factor". To do this, observed data on the actual choices made are normally required: this is known as the Revealed Preference (RP). That observed data are more complex than the clearer-cut comparisons of Stated Preference data making it more difficult to identify the relationships between costs and choice.

1.10.10 Calibration to reproduce the cross-sectional details of the base case is less of an issue when using incremental or pivot-point modelling (the recommended approach), since the observed OD matrices are used directly and only changes from the base or pivoted reference case are to be forecast. In general, Stated Preference methods cannot be relied on for estimating the scales of the responses accurately, and they tend to overestimate the response to change.

1.10.11 Whatever approach is used, a check should still be made that the model reproduces the modal-split correctly for all major flows in the base year in the face of no cost changes, especially when using parameters from other local studies such as a regional Multi-Modal study. In absolute models, adjustment (K) factors may be needed to achieve this. These usually take the form of constants added to the generalised costs to account for the inherent attractiveness of the different modes in the absence of differences in costs. If, for example, a smaller fraction of travellers is observed to use public transport than the estimated difference in generalised cost, and the selected sensitivity (lambda) parameter value (see Section 1.4), suggests, then the addition of a public-transport-specific constant to the public transport cost can adjust the balance to the observed modal split. The constants may vary according to the type of travel modelled, i.e. by purpose, or by region, or by destination.

1.10.12 For the type of scheme assessment considered in this advice, however, it is desirable to keep the number of adjustments of this type to a minimum, even if this means that the modelled base-case departs from the observations in some respects. A great advantage of incremental models is that this type of fitting is avoided, since the model merely predicts the relative changes from the observed base, with all its cross-sectional and category detail, and in a multinomial logit model any mode-specific constants have no effect on the predicted changes. Even so, it is important to ensure that the incremental model does indeed predict no change from the base for no change in costs.

Distribution

1.10.13 Calibration of a trip distribution model can be more difficult and to fit the observed data sufficiently well the calibration may need to be done separately for different sectors. Even if this is done, the estimated parameters are not necessarily the correct values for estimating the responses to changes in costs, since the observed trip patterns occur for a range of historical and land-use reasons not necessarily closely linked to travel costs. To provide a satisfactory local calibration the data available must be of sufficient quality and quantity. This will require that either the range of trip lengths in the observed part of the trip matrix on which the distribution parameter(s) are being calibrated is representative of the whole trip matrix or account is taken of the variations in sampling rate over the full range of trips. The aim is to ensure that the synthesised trip length distribution is correctly representative of the full range of trips.

1.10.14 In practice the main approach to calibrating distribution models is to use observed data.

1.10.15 Given observed or part-observed/part-synthesised trip matrices, it is possible to estimate parameter values based on the present-day distribution of trips quite easily, provided a simple form of distribution model formulation is chosen. Single parameter models (i.e. the lambda of a logit model or the elasticity of a power function) are calibrated by adjusting the parameter iteratively for each calibration area until the average cost (for an exponential or logit function) or the weighted average of the logarithmic costs (for a power function) equals that observed. The theoretical background to the method of calibrating a demand function with a single parameter is available in standard texts such as that by Ortuzar and Willumsen (2001), and in the references for matrix manipulation programs in such transportation modelling suites as Cube (TRIPS), SATURN and EMME/2. A similar approach can be adopted whether a singly or doubly constrained model form is assumed. In practice, the simple demand function may not fit the trip pattern well, and the expected trip-length distribution should be checked against the observed distribution, even where the mean values are well estimated. In addition, the observed trip pattern is likely to contain particular movements that are not properly represented by the modelled function, and additional constants will be needed to reproduce the observed base. As with mode choice use of an incremental model rather than an absolute one avoids most of this complication.

1.10.16 The output values of the sensitivity parameters are then assumed to control the response of travellers' trip distribution to changes in travel costs.

Trip Frequency

1.10.17 If a trip frequency response is included in a variable demand model, the parameters which govern the response to cost changes will be dependent on what other responses are in the model. The 'true' travel cost will be that signified by the composite cost derived from responses lower in the hierarchy: i.e. costs and hence “accessibility” will depend upon the trip distribution and mode-split mechanisms (see Section 1.9). To disentangle these complex interactions unambiguously requires data on responses to large changes in travel costs, and there have been no such studies in this country, although a study of the consequences of completing the Manchester Motorway M60 Box is underway.

1.10.18 In most cases any trip frequency responses to changes in travel costs will be quite small (especially if trips by the slow modes are included).

Time Period Choice

1.10.19 Advice on calibrating models concerning time-period modelling is likely to be modified by the on-going research into better techniques for modelling choice between time-periods, whatever their length. SP techniques can be used to estimate travellers' broad time-period switching in response to travel cost changes, and this may be especially appropriate if one of the policy options relates to encouraging time-period switching. Otherwise, at present, the analyst should follow the advice given in Section 1.11.

General

1.10.20 Calibration of many of the travel responses for specific situations will depend to some extent on SP work, and it will normally be possible to cast the survey methodology to consider more than one demand response, making the use of such techniques more cost-efficient. However, for the general assessments at which this advice is aimed it is likely that such surveys will be not be warranted. Where such sources are not available, and in general where an incremental model is used, the parameter values should be taken from relevant local modelling work or based on the illustrative values of Section 1.11.

1.11 Illustrative Parameter Values

1.11.1 As explained in paragraph 1.9.14, wherever possible, each variable demand response should be calibrated on local data, to reflect the local strengths of the choice mechanisms. Where calibration is not possible, parameter values may be derived from existing, locally calibrated, models of the area, or they may be based on the illustrative values provided in this section. In general, use of values from existing, locally calibrated, models should be considered before adopting the illustrative values given below.

1.11.2 This section suggests illustrative values obtained from a review of a number of UK transport models for situations and responses where either local calibration or derivation from existing models and/or local knowledge is not possible. The values should be compared with local values or modified in the light of local circumstances and accompanied by realism tests. The illustrative values can provide an acceptable approach to including variable demand modelling in transport appraisals where it is deemed too difficult to establish local values.

1.11.3 No matter how carefully the model has been constructed and coded, if the parameter values it contains are wrong the appraisal will be wrong. The base year demand matrix and travel costs will be based on measured local data. It should present a reasonably accurate account of the prevailing situation, but the mechanisms which model travellers' behaviour, and the choices they make, must be calibrated against appropriate evidence of that behaviour. That should ideally include evidence of how choices change as costs change rather than the observed cross-sectional variations.

1.11.4 Although locally calibrated parameters should be used wherever possible, some of the sensitivity parameters may have to be obtained from generalisations of other modelling work. The illustrative values given in the tables below are values obtained for transport models which have been developed by means of rigorous estimation processes. They are not necessarily appropriate for all circumstances, and need to be assessed and modified where necessary but, in the absence of a specific local calibration, they may be the best available estimates. **Whatever values are selected, whether from local knowledge or based on the illustrative values, it is essential to conduct "realism" tests (see *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4)) to ensure that the actual behaviour of the model against variation in travel times and costs accords with experience.**

1.11.5 The illustrative values have been obtained from a review of transport models which have been developed, for areas in the UK, by means of rigorous estimation processes. The parameter values for main mode and destination choice have been derived from "Multi-Modal Model Data Provision", by MVA, dated June 2005. Information was also obtained from Rand Europe on the PRISM model of the West Midlands, but the parameter values in this model are not easily applicable in the models recommended in this guidance due to the fact that the generalised cost formulations and coefficients were determined from local data and do not therefore accord with the advice in Units 3.10.2 and 3.5.6 (see "The PRISM Model: Evidence on Model Hierarchy and Parameter Values" by Charlene Rohr of Rand Europe, dated 3 May 2005). Macro-time period choice parameter values are not provided because the only advice available at this time, from research conducted for the Department, is that these parameter values should be broadly similar to those for main mode choice. The trip frequency parameter values were derived in part from "User-friendly multi-stage modelling advice, Phase 2: Modelling parameters, calibration and validation", by TRL, dated November 2001 – this report has now been largely superseded by the MVA report but may still offer some useful insights. These illustrative parameter values represent the current "best estimates" but are inevitably uncertain.

1.11.6 The Department is keen to obtain further evidence on illustrative values and would welcome information on parameter values from models that have been rigorously calibrated.

1.11.7 All the illustrative parameter values provided in this section relate to generalised costs in minutes, as derived using the Department's standard formulations of generalised cost (see Unit 3.10.2) and standard values of time (see Unit 3.5.6).

1.11.8 If other units or some functional form other than logit were to be used, it is always possible to ensure that the model sensitivity, measured for the local circumstances, is equivalent to that of a logit formulation using the default values as follows:

- estimate "typical" values of the relevant generalised costs;
- apply a modest change to a time or cost component;
- calculate the appropriate change in demand using both the logit formulation and the functional form of the model; and
- adjust the parameters of the model mechanism to obtain a similar change in demand to that given by the logit form.

Trip Frequency

1.11.9 If slow modes are represented in the model, the overall trip rates from each zone can be considered to be constant, and not responsive to changes in travel costs. If slow modes are omitted, then a small sensitivity value can be assumed for the more optional or elastic trip purposes. As *VDM Scope of the Model* (TAG Unit 3.10.2) notes, there should be segmentation by trip purpose as a minimum, and some purpose categories, such as work and employer's business, should be regarded as inelastic. Note that here we are concerned solely with the response of the total number of trips from each zone changing as travel costs change: dependence on the demographics and land use of the zone is a different issue. If the demand model includes slow modes, then trip rates can be assumed not to be responsive to changes in travel cost.

1.11.10 Some models include trip frequency but the evidence on the appropriate sensitivity parameter value is limited and as a result we do not currently have any suitable recommended values.

Destination Choice

1.11.11 Illustrative destination choice parameter values are shown in the table below. On the presumption that destination choice will follow main mode choice in the model hierarchy (see Section 1.9), parameter values are provided separately for car trips and public transport trips. See Appendix 4 for the model formulation to which these parameter values apply.

1.11.12 The parameter values for public transport trip distribution will only be required in fully-specified multi-modal models – that is, they will not usually be required for a model designed solely for the appraisal of highway schemes.

TRIP PURPOSE AND MODE	MINIMUM	MEDIAN	MAXIMUM	SAMPLE
CAR				
Home-based work	0.054	0.065	0.113	7
Home-based employers business	0.038	0.067	0.106	5
Home-based other	0.074	0.090	0.160	4
Non-home-based employers business	0.069	0.081	0.107	3
Non-home-based other	0.073	0.077	0.105	3
TRIP PURPOSE AND MODE	**MINIMUM**	**MEDIAN**	**MAXIMUM**	
PUBLIC TRANSPORT				
Home-based work	0.023	0.033	0.043	7
Home-based employers business	0.030	0.036	0.044	4
Home-based other	0.033	0.036	0.062	4
Non-home-based employers business	0.038	0.042	0.045	2
Non-home-based other	0.032	0.033	0.035	3

1.11.13 The parameter values shown above for public transport trips strictly apply to trips from car-available households. They may also be used for trips from non-car available households without significant loss of accuracy.

1.11.14 It is difficult to generalise about when low values should be used and when high values would be more appropriate. (Note that the ranges shown above are not targets within which parameter values must lie; they are simply the minimum and maximum values from the sample available.) The MVA Report provides parameter values for a variety of models, of London, a large region in Scotland, and a number of smaller urban areas. This report should be consulted in deducing parameter values for models of more complex areas where the use of the single mean or median values may be considered too simplistic. The TRL Report may also provide some guidance on variations in parameters under different circumstances, although it should be borne in mind that this report contains limited information about the extent to which model parameters were derived by rigorous calibration procedures and validated by realism tests.

Main Mode Choice

1.11.15 Main mode choice (that is, the choice between car and public transport) parameters are specified as scaling parameters –see Appendix 4 for further details. These scaling parameters show the sensitivity of main mode choice relative to destination choice. Thus, to be consistent with the default hierarchy recommended in Section 1.9, of destination choice following main mode choice, the main mode choice scaling parameters are all less than or equal to one, as shown in the table below.

TRIP PURPOSE	MINIMUM	MEDIAN	MAXIMUM	SAMPLE
Home-based work	0.50	0.68	0.83	6
Home-based employers business	0.26	0.45	0.65	2

TRIP PURPOSE	MINIMUM	MEDIAN	MAXIMUM	SAMPLE
Home-based other	0.27	0.53	1.00	4
Non-home-based employers business	0.73	0.73	0.73	1
Non-home-based other	0.62	0.81	1.00	2

1.11.16 Again, it is difficult to generalise about when low values should be used and when high values would be more appropriate. (Note that the ranges shown above are not targets within which parameter values must lie; they are simply the minimum and maximum values from the sample available.) The MVA Report should be consulted in deducing parameter values for models of more complex areas where the use of the single mean or median values may be considered too simplistic. The TRL Report may again be of some use, noting the caveat in paragraph 1.11.13.

Time of Day Choice

1.11.17 Less evidence is available about the sensitivity of the macro-time period choice than either main mode or destination choice. Recent research conducted for the Department suggests that the sensitivity of the choice between relatively long time periods, such as three hours or so, should be about the same as that of main mode choice. The research also suggests that, as the time periods are reduced, the sensitivity increases. Thus, when long time periods, of the order of three hours, are being modelled, macro-time period choice should be positioned either just before or just after main mode choice, with parameter values similar in magnitude to the main mode choice parameter values. Peak spreading, or micro-time period choice that will include a schedule disutility term, if included in the model, should be positioned after destination choice.

Testing

1.11.18 Whatever values are selected, whether from local knowledge or based on the illustrative values, it is essential to conduct “realism” tests (see *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4)) to ensure that the actual behaviour of the model against variation in travel times and costs accords with experience.

2 Further Information

The following documents provide information that follows on directly from the key topics covered in this TAG Unit.

For information on:	See:	Link:
General Multi-modal models	*TAG Unit Introduction to modelling*	TAG Unit 3.1.1
An overview of modelling issues	*Summary Advice on Modelling*	TAG Unit 2.4
An overview of variable demand modelling	*Variable Demand Modelling*	TAG Unit 2.8
Detailed advice on transport modelling	*Modelling*	TAG Unit 3.1

3 References

DETR (July 1998) *A New Deal for Transport: Better for Everyone*

DfT (2003) Transport Economics Note (TEN)

Ortuzar, J de D and Willumsen, L G (2001). Modelling Transport (Third Edition). Wiley, Chichester

Senior M L and H C W L Williams (1977) Model-based transport policy assessment. *Traffic Engineering and Control*, Vol 18 (9/10).

Institution of Highways and Transportation (1996) *Guidelines on Developing Urban Transport Strategies.*

4 Document Provenance

This Transport Analysis Guidance (TAG) Unit reflects the consultation comments received on the Key Processes Stage of the draft Variable Demand Modelling Advice produced by TRL in June 2003.

Section 1 of this document has been modified in October 2009 following the release of Proportionate Appraisal Guidance.

Technical queries and comments on this TAG Unit should be referred to:

Integrated Transport Economics and Appraisal (ITEA) Division

Department for Transport

Zone 3/04 Great Minster House

76 Marsham Street

London

SW1P 4DR

itea@dft.gsi.gov.uk

Tel 020 7944 6176

Fax 020 7944 2198

1 Appendix 1 Elasticity Models

1.1 Functional Forms and Parameter Values of Elasticity Models

1.1.1 Where an elasticity model is appropriate the functional form and parameter values need to be selected. The simplest functional form – an ‘own-cost’ elasticity model - assumes that changes in the demand for travel between two points can be adequately estimated purely by a function of the change in costs between the two places.

1.1.2 **‘Own-cost’ elasticity models** assume that the demand for travel between two points is purely a function of the change in costs on that mode between the two places. The strength of that function can vary for different trip lengths.

1.1.3 However, if costs do indeed change, the relationship between change in demand and change in costs can take a number of forms, but only exponential and power formulations, and a composite of the two forms (called a Tanner function), will be considered here. With a power formulation the proportionate change in trips is related to the proportionate change in costs, as shown in the equation below. With an exponential formulation, on the other hand, the proportionate change in trips is a function of the absolute change in costs. Other, more complex, relationships are described in *VDM Appendices* (TAG Unit 3.10.5).

1.1.4 For most applications the Power relationship below which is a simple own cost elasticity model due to its constant elasticity value is recommended:

$$T_{ij} = g_{ij} * {}_0T_{ij} * (G_{ij} / {}_0G_{ij})^{A}$$

Where 1. T_{ij} is the forecast number of trips between zones i and j

2. G_{ij} is the forecast disutility or generalised cost

3. g_{ij} is the forecast growth rate relative to an earlier or base year

4. ${}_0T_{ij}$ is the number of trips in the earlier or base year

5. ${}_0G_{ij}$ is the disutility or generalised cost in the earlier or base year

6. A is the elasticity, which should be negative and is the same for all trips in the same user class.

1.1.5 This is a well-behaved formulation that is simple to apply, and is base independent: that is, it is guaranteed to give the same results if forecasts are produced from one year to another directly or via an intermediate year. It assumes that a *proportionate* change in trips is related to a proportionate change in costs. As the parameter A is constant the implied elasticity is the same for all lengths of trip within the same user class (i.e. it is “distance neutral”).

1.1.6 This formulation can easily be set up using the matrix manipulation facilities available in modern transportation modelling suites, or in some modelling suites combined directly with the assignment process. The facility is also available within the DIADEM modelling framework.

1.1.7 An alternative formulation is the “Exponential” relationship. In this case the effective elasticity increases with increasing trip cost, and hence for study areas where there are a wide variety of trip lengths the effective elasticites could vary markedly. Thus the exponential approach should only be considered in the case where the study area is small and urban, and where a general elasticity approach is being combined with a logit choice mechanism to jointly represent the individual demand mechanisms. Most logit mechanisms in the variable demand hierarchy share this exponential function characteristic, but some have a more benign effect since trip re-distribution, for example, can be constrained to avoid changing the overall number of trips. In that case trip re-distribution in the face of changing travel costs effectively adjusts the proportions of trips of different length to compensate for the changes. Similarly mode-choice models estimate shares rather than absolute numbers.

1.1.8 The "Exponential" formulation assumes that the proportional change in trips is a function of the absolute change in costs:-

$$T_{ij} = g_{ij} * {}_0T_{ij} * \exp\{B*(G_{ij} - {}_0G_{ij})\}$$

Where the elasticity of demand with respect to generalised cost U is $B*G_{ij}$ with B negative. This is an own cost elasticity which is not simple due to the elasticity not being constant.

1.1.9 These equations can be used in two ways. They can be based or pivoted on a base year, where items 4 and 5 in paragraph 1.1.4 (base trips and costs) are known from empirical data, and the product $g_{ij} * {}_0T_{ij}$ represents what is referred to as the Reference Case Matrix, see *VDM Scope of the Model* (TAG Unit 3.10.2). Alternatively, the equations can be formulated to compare costs between alternatives for the same year, where the 'earlier' year costs and trips are derived from the other scenario. More details on how these equations are to be used are given in *VDM Appendices* (TAG Unit 3.10.5). This should be read in conjunction with the guidance given with the software employed to undertake the transport modelling (see also *VDM Convergence Realism and Sensitivity* (TAG Unit 3.10.4).

1.1.10 Both these formulations are closely incremental in nature, allowing the number of trips in the system to change up or down. This is in contrast to most of the individual demand-response mechanisms that are set out as share formulations where the total number of trips is fixed (say by TEMPRO all-mode forecasts) and merely allocated to one choice or another (e.g. to different modes or destinations).

1.1.11 Whilst the formulation is relatively easy to set up, there are some issues that must be dealt with when considering the parameter values to assign to a demand segment.

1.1.12 'Own cost' elastic assignment modelling in congested Urban Areas should be undertaken at a peak hour unless there are significant variations in demand, or congestions levels are high in which case the modelling should cover linked time-periods, sub-divided into time slices and sub-periods.

1.1.13 The size of the parameter value will reflect the number of responses that the elasticity formulation is acting as proxy for. For instance, if the elasticity formulation is taking the place of all responses then it will be larger than if it is acting as proxy for only one or two responses. The table below sets out the recommended starting values for the elasticity of demand with respect to journey time.

Table 1 Derived long-term journey time elasticites for different uses (derived from 1997 DMRB Vol 12 Section 2 Part 2 Table C2 and its para C13)

Purpose	Time elasticity – High modal competition	Time elasticity – Low modal competition	Time elasticity – High modal competition including time-switching	Time elasticity – Low modal competition including time switching	Trip frequency effect (only)
HB Work	-0.22	-0.14	-0.48	-0.30	-0.04
Employer's	-0.60	-0.35	-0.96	-0.55	-0.15

Business					
Essential Other	-0.47	-0.26	-0.65	-0.36	-0.12
Discretionary Other	-0.35	-0.20	-0.50	-0.28	-0.10
Note:- The values are based on car journey time elasticities - equivalent generalised cost elasticities would be about 10-50% higher, depending on the value of time and average network speed. Short-term elasticities are 28%, 8% & 5% less for HBW, Employer's Business and Discretionary purposes.					

1.1.14 Equivalent journey cost elasticities can be calculated from the above table by dividing the elasticities by the proportion of the total generalised cost made up of journey time. For instance, if a model assigns on the basis of generalised cost (t+kd), the appropriate time elasticity must be multiplied by a factor (1+kv) where v is the average speed in the base year in kilometres per minute if journey time (t) is in minutes and distance (d) is in kilometres. *Values of Time and Operating Costs* (TAG Unit 3.5.6), can be used to provide the relevant factors for given combinations of purpose, forecast period and congestion level if standard values of time are being used. In practice, the generalised cost elasticities will be between 10% and 50% higher than the values shown in the table above with values at the lower end for Employer's Business trips, urban areas and later forecast years.

1.1.15 If an exponential formulation is used then the above values will need to be subsequently divided by the mean generalised cost to give the equivalent parameter value.

1.1.16 The estimated generalised cost elasticities (and associated parameter values if an exponential model is used) may need to be adjusted so that the fuel cost elasticity estimate from the model reflects the national overall estimate of -0.3 (see *Variable Demand Modelling – Convergence Realism and Sensitivity* (TAG Unit 3.10.4)).

1.1.17 **Where possible, the trips should be split by trip purpose (and any other known major variation such as willingness to pay or movement type). If this is not possible, for instance where only a single private vehicle user class is available, then they should be split by time-period.** *VDM Appendices* (TAG Unit 3.10.5) shows how, by using the national car driver journey purpose mix for each period of the day (from NTS), the above elasticities can be converted to elasticities for all trips by time period. If local data suggests a significantly different mix of purposes by time-period, then the local proportions can be substituted for the national ones.

1.1.18 Care should be taken when dealing with intra-zonal trips. Because most assignment models do not output intra-zonal costs (since intra-zonal trips are not assigned) there may be problems with using incremental models where there are observed intra-zonal trips in the base year trip matrix. It is desirable that robust estimates of intra-zonal costs should be estimated in these instances. These could be some function of the inter-zonal costs, for example half the minimum inter-zonal costs for that zone. Further advice is given in section *VDM Scope of the Model* (TAG Unit 3.10.2). Power function elasticity models will be particularly sensitive to very small intra-zonal costs, and this is one reason why they should be avoided when this is the case

2 Appendix 2 Functional Forms for VDM

2.1 Detailed Advice on Functional Forms of VDM

2.1.1 There are various mathematical functions that can provide a suitable relationship between travel demand and the disutility or generalised cost of a trip. These all offer broadly similar behaviour, but have subtly different mathematical properties. Section 1.1 of Appendix 1 discussed the equivalent subtle differences of power functions and exponential functions for own cost elasticity models, both of which can provide a convenient downward sloping relationship as shown in Figure 1 in Section 1.3 and the parameters can be adjusted to give an elasticity of any required strength.

2.1.2 Most of the mechanisms required in a variable demand model allocate trips between a set of choices, giving rise to the slightly more complex relationship of Figure 2 in Section 1.3. Here again a range of mathematical functions, most based on powers or exponentials, or both, can recreate the desired relationship. *VDM Appendices* (TAG Unit 3.10.5) discusses the detailed functional forms of VDM models and the derivation of many of the forms, while for example Ortuzar and Willumsen (2001) provides even more detail.

2.1.3 The multinomial logit choice function is one of a number of possible formulations of "random utility" models in which a random component is added to the deterministic utility of choice p as follows:

$$U_p = \sum_n \beta_n x_n + \varepsilon_p$$

Where the utility (or disutility) U_p of choice p is calculated as the sum of:

- the generalised cost $G_p = \sum_n \beta_n x_n$ of choice p, with the set of cost components x_n weighted by coefficients β_n, summed over all components relevant to choice p as explained in *VDM Scope of the Model* (TAG Unit 3.10.2). (For example, if G is measured in units of time, x might be the money cost of a journey and β the inverse of Value of Time), and
- a random component ε_p used to represent variations in the situation or tastes of individual travellers, or modelling errors, or unobserved elements of the alternative choices. (In the most general case this random component can depend on both the traveller and on the choice alternative).

A choice-specific calibration constant (i.e. constant specific to the mode or areas used in calibration) could be added to the generalised cost function to adjust the calculated choice to the observed value.

A random utility model assumes that the alternative with the maximum utility (or minimum disutility) is chosen, so that a probabilistic model results.

2.1.4 The assumed statistical distribution of the error terms or residuals ε_p determines the exact mathematical formulation. For example, assuming one particular distribution for the random components, that they are Independent and Identically Distributed (IID) extreme value variables, leads to the widely-used multinomial logit model:

$$T_p = T_{tot} \exp(-\lambda U_p)/\{\sum_q \exp(-\lambda U_q)\}$$

2.1.5 Conventionally, different Greek symbols have been used for this sensitivity parameter λ according to the mechanism it is applied to (for example α for trip frequency, and β for mode split), but usage varies and here we will use the λ formulation for all applications, distinguishing between the different mechanisms of variable demand by a subscript.

2.1.6 The elasticity of demand in this formulation is $-\lambda U_p(1-T_p/T_{tot})$, so that the elasticity scales with U, and tends to be larger for longer trips for a given value of λ and larger for choices with a small share of the total. If those implications are inappropriate for the model area a different functional form or a series of calibration areas should be used to produce a model with suitable implications.

2.1.7 Other forms such as the power function or the Tanner function, which have been described in relation to 'own cost' elasticity models, or formulations assuming a normal distribution of error terms (*Probit* models) are possible but little used in modelling for scheme appraisal. However, different formulations of the logit model which have less restrictive statistical assumptions are also possible and are being investigated in current research (See *VDM Appendices* (TAG Unit 3.10.5) for more details).

2.1.8 The remainder of this Advice focuses on hierarchical and multinomial logit. **The logit formulation (and its nested variants) can be used, in slightly different formulations, for each of the mechanisms of the variable demand model. These are discussed in turn below.**

Trip Generation and Frequency

2.1.9 In the trip generation stage the numbers of trips of different kinds made from each zone can be predicted as a function of the numbers of different types of households and inhabitants in the zone. For the purposes of this advice, however, we are interested primarily in how the externally estimated demand from each zone might respond if travel costs change, i.e. trip frequency change. The growth factors derived for fixed OD matrices usually assume that travel costs would be unchanged from the base situation. If so they are known as a reference case and are a suitable starting point for variable demand modelling. If as part of that modelling trip generation is to be made responsive to changes in travel cost (see Section 1.4) the requirement is to make trips in any category elastic to changes in cost and thus model trip frequency.

2.1.10 The elasticity function could be a power function or an exponential function as described in Section 5.1. However, if logit is used for the other mechanisms, a similar exponential function is generally used to adjust trip frequency. In this case, the function operates simply as an elasticity with respect to disutility or generalised cost, since the relevant choice is to travel or not to travel, and the disutility of not travelling remains constant:

$$T_i = {}_0T_i \exp(-\lambda_{freq}\, G_{icomp})/\exp(-\lambda_{freq}\ {}_0G_{icomp})$$

Where T_i is the number of trips from origin zone i, prefix 0 denotes the base values, λ_{freq} is the choice sensitivity parameter for the trip frequency stage and the generalised cost G_{icomp} is the composite cost or disutility calculated across the trip origins.

Composite Costs

2.1.11 In the above and equivalent equations the disutility (or generalised cost) G_{icomp} must be calculated to represent the "average" perceived cost, or compound cost, across all alternative choices available at lower levels in the hierarchy. Thus for trip frequency the compound cost of travel for trips from a given zone must be calculated across all available choices of destination zone, mode and time-period if the latter is included. G_{icomp} is intended to provide an estimate of the likely average cost from zone i and incorporates the probability of making each choice, to give a "logsum" cost as described in Section 1.9. This calculation is applied to each category of travel (e.g. trip purpose by SEG) in each origin zone separately. When a scheme is introduced, only trips from those zones where an appreciable proportion of trips experience the scheme, or its surrounding effects, will be noticeably affected.

Trip Distribution

2.1.12 Distribution models spread the generated trips over the available destinations, depending on the generalised cost of reaching that destination. Early demand modelling based distribution on distance rather than cost, and often as a simple negative power function of distance. These were known as "gravity" models in analogy with the gravitational attraction between masses, but use of logit functions based on utilities or generalised costs is now almost universal in demand models. In addition to cost, distribution also depends on some measure of the attraction of a zone, estimated in terms of the numbers of "opportunities" such as jobs or retail floorspace in the zone.

2.1.13 Most distribution models are designed to guarantee that the total number of trips from the origin zone (or to the destination zone) is equal to the total number of trips for that zone forecast at the trip generation/frequency stage. If the design guarantees that property for both origins and destinations the model is known as doubly constrained.

2.1.14 Some distribution models do not depend on travel costs and merely estimate future OD matrices directly from a base-year matrix and the future row and column totals. These methods are known as *Fratar* or *Furnessing* and are used to provide reference case growth factors for movements between zones to reflect the forecast growth in zonal trip ends. Apart from that role in providing inputs to variable demand modelling they are not relevant in the context of multi-stage demand modelling.

2.1.15 The general form for a doubly-constrained distribution model is:

$$T_{ij} = a_i * b_j * O_i * D_j * f(G_{ij})$$

Where T_{ij} is the number of trips from zone i to zone j,

O_i is the total number of trips originating in zone i

D_j is the total number of trips ending in zone j

T_{ij} depends on the travel disutility or cost via the deterrence function as $f(G_{ij})$,

Which in most models is a logit function

$$\exp(-\lambda_{dist}\ G_{ijcomp})/\{\textstyle\sum_k \exp(-\lambda_{dist}\ G_{ikcomp}),$$

Where G_{ijcomp} is a composite cost calculated across the available modes and time periods, if these choices are to be calculated *after* distribution

a_i and b_j are balancing factors which are only used when the model is singly or doubly constrained (see Section 1.17) to ensure that $\sum_j T_{ij} = O_i$ (ie there are O_i trips originating in zone i), and $\sum_i T_{ij} = D_j$ (ie there are D_j trips ending in zone j), and are calculated at each iteration of the constraining routine as $a_i = 1/\sum j\ b_j\ D_j\ f(G_{ij})$ or $b_j = 1/\sum i\ a_I\ O_I\ f(G_{ij})$

(In the above equation, T_{ij} is proportional to O_i, the total number of trips originating in zone i,T_{ij} is also proportional to D_j, the total number of trips ending in zone j, or alternatively in a singly-constrained model the facilities available in zone j (numbers of jobs, retail floorspace, etc) so that the number of trips ending in j depends also on the competing attractivities of other zones. (See VDM Scope of the Model (TAG Unit 3.10.2) *for conversion from Production/Attraction matrices to O/Ds)).*

Mode Choice

2.1.16 For mode choice, trips between each origin-destination pair of zones are allocated to the available modes according to the composite disutility or generalised cost of travel by that mode:

$$T_{ijn} = T_{ij} \exp(-\lambda_{mode}\ G_{ijn})/\{\sum_m \exp(-\lambda_{mode}\ G_{ijm})\}$$

if mode choice is the only demand response; and

$$T_{ijn} = T_{ij} \exp(-\lambda_{mode}\ G_{ijncomp})/\{\sum_m \exp(-\lambda_{mode}\ G_{ijmcomp})\}$$

if mode choice is a more sensitive response than distribution. T_{ijn} is the number of trips choosing mode n from a set of modes m and λ_{mode} is the choice sensitivity parameter for the trip mode stage. The composite cost $G_{ijncomp}$ is calculated across the time periods in a way that weights the average according to the probability of choosing that period. The summation is across all available modes m, including n. However, if mode choice is less sensitive than distribution, the composite cost G_{imcomp} must be calculated to forecast an overall modal split for each origin zone. If there is more than one public transport mode it is usual to use a nested or hierarchical model, with a higher level split between car and public transport (and possibly slow modes also). The allocation to the different public transport modes (and between walk and cycle if modelled) is then made at a lower level (see Section 1.9) or possibly in assignment.

Time of day choice

2.1.17 **Macro time period choice** (or the allocation of trips between broad time periods) assuming this is the most sensitive response takes the form:

$$T_{ijms} = T_{ijm} \exp(-\lambda_{time}\ G_{ijms})/\{\sum_t \exp(-\lambda_{time}\ G_{ijmt})\}$$

Where T_{ijms} is the number of trips between zones i and j by mode m in time period s. G_{ijmt} is the disutility or generalised cost of travel between zones i and j by mode m in time period t, which may typically be peak and inter-peak and λ_{time} is the choice sensitivity parameter for the time period stage. However, if it is above mode choice and distribution, it would take the form:

$$T_{is} = T_i \exp(-\lambda_{time}\ G_{is})/\{\sum_t \exp(-\lambda_{time}\ G_{it})\}$$

Where T_{is} is the number of trips in Zone i in time period s and G_{is} is the disutility or generalised cost of travel in zone i in time period s.

2.1.18 It should be noted that the sequence of responses given in this section is arbitrary and should not be taken as the recommended structure.

2.1.19 Research is underway into the modelling of **micro time period choice** to improve the robustness of current models. Research results can be found at www.dft.gov.uk/stellent/groups/dft_econappr/documents/divisionhomepage/032182.hcsp.

3 Appendix 3 Example Estimation of Modal Split

3.1.1 Modal split for those travellers who have a car available is likely to be substantially different from the overall split across all travellers. The method described below is very approximate, but it gives a general indication of how important the alternative public transport is likely to be (or for any alternative mode, since the principle is the same).

3.1.2 First, it is necessary to estimate the generalised costs by the car and non-car modes (bus in this example) for trips affected by the scheme being assessed. This may involve several different groups of traffic movements, in which case the estimate should be made for an “average” journey in each group (though, for the purposes of this exercise, the judgement of an "average" or "typical" journey can be very approximate). The generalised costs are as follows

Car Travel: G = 2*A + 60*D/V + D*VOC/(occ*VOT) + PC/(occ*VOT)

where A is the access time at both ends of the trip: since this will generally be walked, it is weighted by 2; D is the mean journey distance in kms; V is the mean traffic speed in kph; occ is the mean car occupancy; VOC is the mean car operating cost in pence per km; and VOT is the value of time per person in pence per minute; PC is half the mean car parking cost in pence. G is measured in *minutes* in the calculations described here.

Bus travel: G = 2*Walk + 2*Wait + 60*D/V + Fare/VOT + I

Where Walk is the sum of the time spent walking to the stop or station at the origin end of the journey and the time spent from the destination stop or station to the actual destination, and Wait is the mean time spent waiting for the service, which will be half the service interval for frequencies of two per hour or better, rising to a nominal 15 (minutes for less frequent services. Walk and Wait are weighted by 2 since this time is costed highly. V is the mean journey speed, including stopping. The mean fare appropriate to those travellers likely to choose between car and public transport (i.e. probably not concessions) should be for a single journey divided by the mode-specific value of time as for car. If access is by Park and Ride it will be necessary to add in the car access generalised cost, including half any parking fee, and if access at either end of a rail journey is by bus the extra generalised cost should be calculated as described here and added on. If the journey requires interchange between services the extra wait time (and walk time if relevant) should be included and an interchange penalty of, say, 6 minutes per interchange added.

3.1.3 Taking as an example the following cost components:

	Car trip	Bus trip
Access walk – both ends (mins)	3	10
Wait time (mins)	NA	8
In-vehicle time (mins)	30	50
Fuel, fares and other costs (pence)	160	250
Park cost - half of charge per one-way trip(pence)	0	NA
Interchange penalty (mins)	NA	5
Mean occupancy	1.3	NA
Generalised cost (mins)	51.4	122.2

Then the generalised cost in the last line is obtained via the calculations above, using a Value of Time of 11 pence/minute.

3.1.4 The notional modal split is calculated as:

$$P_{PT} = \exp(-\lambda_{mode}G_{PT})/\{\exp(-\lambda_{mode}G_{PT}) + \exp(-\lambda_{mode}G_{car})\}$$

Where P_{PT} is the proportion of travellers choosing public transport, G_{PT} and G_{car} are the generalised costs of travel by public transport and car respectively, and λ_{mode} is the mode choice sensitivity parameter, (which can be given a value of say 0.04), unless there is local knowledge of the prevailing value. This gives 5.5% to bus, but note that this calculation **does not include any mode-specific constant** beyond that implied by any mode specific value of time: it is not possible to generalise about the values of these constants, but they are generally found to reduce the share to public transport, so the 5.5% estimated in this case is likely to **over-estimate the use of bus by car-available travellers.**

3.1.5 Consider a highway improvement that saves 1.5 minutes in journey time for road traffic. The bus has to stop periodically and with acceleration and deceleration cannot make full use of the higher road speeds. Assume it gains only half this time, 0.75 minutes. Then in the "after" case, the car generalised cost falls to 49.9 minutes, and the bus generalised cost to 121.5 minutes. The mode split to bus falls to 5.4% whereas, if it had been assumed that the bus journey time had not changed, the reduction in car generalised cost would reduce the mode split to bus to 5.2%. Thus in this example inclusion of the effect on bus times means that the modal share hardly changes from the "before" situation.

3.1.6 However, although the effect on mode split is very small, this gain in journey time for bus travellers may account for an appreciable part of the total economic benefit. Car-available travellers account for only 5.4% of the total flow, and with only half the saving in travel time they account for only 2.7% of the benefit to car users. However, bus users who do not have a car available also gain this benefit, and they are likely to be much greater in number than the car-available bus users.

3.1.7 Such detailed assessment of the impact on public transport, in the absence of a full public transport modelling will be largely confined to circumstances where no public transport alternative is being proposed but the impacts on public transport of the scheme (or the congestion in the without scheme scenario) are thought to be significant.

4 Appendix 4 Incremental Model Formulation

4.1.1 When specifying an incremental hierarchical logit model, scaling parameters as provided in section 1.11.15 could be used. These parameters refer to the probability of nests of alternatives or composite alternatives. They reflect the ratios of the lambdas for different response mechanisms as you move up the model structure. The scaling parameters are applied to the logsums of the composite or nested alternatives. They should have a value between 0 and 1 if the responses have been included in the correct order in the model, such that the sensitivity of the responses changes down the hierarchy from lower to higher.

4.1.2 The standard incremental multinomial logit model is given as

$$p_i = \frac{p_i^0 \exp(\theta \Delta U_i)}{\sum_j p_j^0 \exp(\theta \Delta U_j)}$$

where

p_i is the forecast probability of choosing alternative i

p_i^0 is the reference case probability of choosing alternative i (calculated from the input reference demand)

θ is the scaling parameter (always =1 for the bottom level of the hierarchy)

ΔU_i is the change in the utility of alternative i

For the choice at the bottom level of the hierarchy the change in utility is given by

$$\Delta U_i = \lambda (C_i - C_i^0)$$

where

C_i^0 is the reference generalised cost and

C_i	is the forecast generalised cost, skimmed from the latest assignment
λ	is the spread or dispersion parameter (defined by the user); it should be negative

For choices above the bottom level of the hierarchy the change in utility is the composite change over the alternatives in the level below:

$$\Delta U^* = \ln \sum_i p_i^0 \exp(\Delta U_i)$$

4.1.3 This model formulation can be used for mode choice, time period choice and singly constrained distribution

A modified version of the logit model is used for doubly-constrained distribution as follows:

$$T_{ij} = O_i \frac{B_j T_{ij}^0 \exp(\theta \Delta U_{ij})}{\sum_{k=1}^{N} B_k T_{ik}^0 \exp(\theta \Delta U_{ik})}$$

where

T_{ij}	is the forecast number of trips travelling from zone i to zone j
T_{ij}^0	is the reference case number of trips travelling from zone i to zone j
O_i	is the number of trips travelling from zone i
B_j	are destination-based constants, normalised so that $\sum_j B_j$ is equal to the number of zones

Note that destination constraints are summed over all person types within a purpose, and across all modes and time periods, if those choices have been modelled.

The change in composite utility for origin zone a is calculated using:

$$\Delta U_a^* = \ln \sum_b B_b \frac{T_{ab}^0}{O_a^0} \exp(\theta \Delta U_{ab})$$

4.1.4 The illustrative parameter values currently provided in Section 1.11 can be used in an incremental model structure as follows:

Suppose we assume the follow choices available

- Single trip purpose (say commuting) split into:
- Two person types (say car available and car not available)
- Car available hierarchy (from top to bottom): frequency, mode choice, time period choice, distribution (doubly constrained)
- Car not available hierarchy (from top to bottom): frequency, time period choice, distribution (doubly constrained)

Inputs

4.1.5 Inputs to the demand model are:

C^0_{ijmtpc} reference generalised cost from zone i to zone j by mode m in time period t, trip purpose p, person type c

C_{ijmtpc} corresponding forecast generalised cost, skimmed from latest assignment

T^0_{ijmtpc} corresponding reference demand, defined via the user interface

In all the above, there is no data for the highway mode for the no-car person type

Bottom level utilities

4.1.6 The first step is to calculate the change in utility for the lowest level of the hierarchy:

$$\Delta U_{ijmtpc} = \lambda^{mode}_{mc}\left(C_{ijmtpc} - C^0_{ijmtpc}\right)$$

Where λ^{mode}_{mc} is the mode-specific distribution λ parameter

Doubly-constrained distribution

4.1.7 Since the lowest level is a doubly constrained distribution model we need to find the balancing factors B_{jp} . This requires solving the set of equations given by:

$$T_{ijmtpc} = O_{imtpc}\frac{B_{jp}T^0_{ijmtpc}\exp\left(\Delta U_{ijmtpc}\right)}{\sum_{k=1}^{N} B_{kp}T^0_{ikmtpc}\exp\left(\Delta U_{ikmtpc}\right)}$$

such that the destination trip end constraints are met:

$$\sum_{imtc} T_{ijmtpc} = D_{jp}$$

The destination constraints are calculated from the reference demand matrix:

$$D_{jp} = \sum_{imtc} T^0_{ijmtpc}$$

Note that the destination trip end constraints depend on destination and trip purpose only.

The balancing factors are normalised so that

$$\sum_{j} B_{jp} = N$$

where N is the number of destination zones.

On the first iteration only of the demand model the origin trip ends are calculated from the reference demand matrix:

$$O_{imtpc} = \sum_{j} T^0_{ijmtpc}$$

For subsequent iterations they are obtained from the application of the conditional probabilities described below.

Composite Utilities

4.1.8 The change in the composite utility from the distribution, time period choice and mode choice stages is then calculated:

$$\Delta U^{*}_{imtpc} = \ln \sum_{j} B_{jp} \frac{T^{0}_{ijmtpc}}{O^{0}_{imtpc}} \exp\left(\Delta U_{ijmtpc}\right)$$

$$\Delta U^{*}_{impc} = \ln \sum_{t} p^{0}_{t|\,mcp} \exp\left(\theta^{time}_{c} \Delta U^{*}_{imtpc}\right)$$

$$\Delta U^{*}_{ipc} = \ln \sum_{m} p^{0}_{m|\,cp} \exp\left(\theta^{mode}_{c} \Delta U^{*}_{impc}\right)$$

(car available person type)

$$\Delta U^{*}_{ipc} = \Delta U^{*}_{impc}$$

(car not available person type, m= PT)

The reference case probabilities are calculated from the input reference demand as follows:

$$p^{0}_{m|\,cp} = \frac{\sum_{jt} T^{0}_{ijmtpc}}{\sum_{jtk} T^{0}_{ijktpc}}$$

$$p^{0}_{t|imcp} = \frac{\sum_{j} T^{0}_{ijmtpc}}{\sum_{k} T^{0}_{ijmkpc}}$$

Conditional Probabilities

4.1.9 Having calculated the change in the composite utilities it is possible to calculate the conditional utilities for each level of the model

Mode choice:

$$p_{m|\,pc} = \frac{p^{0}_{m|ipc} \exp\left(\theta^{mode}_{c} \Delta U^{*}_{impc}\right)}{\sum_{k} p^{0}_{k|ipc} \exp\left(\theta^{mode}_{c} \Delta U^{*}_{ikpc}\right)}$$

(car available person type)

$$p_{m|\,pc} = \begin{cases} 1 & if\ \ m = public\ transport \\ 0 & otherwise \end{cases}$$

(car not available person type)

Time period choice:

$$p_{t|impc} = \frac{p^0_{t|impc} \exp\left(\theta_c^{time} \Delta U^*_{imtpc}\right)}{\sum_k p^0_{k|impc} \exp\left(\theta_c^{time} \Delta U^*_{imkpc}\right)}$$

Distribution (destination choice):

$$p_{j|\,mtpc} = \frac{B_{jp} T^0_{ijmtpc} \exp\left(\Delta U_{ijmtpc}\right)}{\sum_{k=1}^{N} B_{kp} T^0_{ikmtpc} \exp\left(\Delta U_{ikmtpc}\right)}$$

Updated Trip Matrix

4.1.10 The application of the conditional probabilities gives an updated trip matrix

$$T_{ijmtpc} = T^0_{ipc}\, p_{m|\,pc}\, p_{t|\,mpc}\, p_{j|\,mtpc}$$

and updated origin totals:

$$O_{imtpc} = T^0_{ipc}\, p_{m|\,pc}\, p_{t|\,mpc}$$

Application of Frequency Model

4.1.11 The frequency model is only applied after the above process has converged. This gives the final trip matrix from the demand model:

$$T_{ijmtpc} = \exp\left(\theta_c^{freq} \Delta U^*_{ipc}\right) T^0_{ipc}\, p_{m|ipc}\, p_{t|impc}\, p_{j|imtpc}$$

5 Appendix 5 Absolute Model Formulation

5.1.1 The illustrative parameter values currently provided in Section 1.11 can be used in an absolute model structure as follows:

Assumed Nesting

Layer 1 (Highest):	Frequency
Layer 2:	Main Mode
Layer 3:	Macro Time Period
Layer 4 (Lowest):	Destination

Notation

Trip Origin:	i	Trip Destination	j, k
Macro Time period	t, s	Main Mode	m, r
Trips	T	Generalised Cost	G
Distribution parameter	λ_{dist}	Attraction Factor	B
Composite Utility	U	Tree parameters	θ_{time}, θ_{mode}, θ_{freq}

Pivot (reference) Trips	${}_0T$	Pivot (reference) Utilities	${}_0U$

Composite Utilities:

5.1.2 The incremental composite utilities summed over the choices in the destination layer are given by:

$$U_{imt} = \ln \sum_j B_j \exp(-\lambda_{dist} G_{ijmt}) - \ln \sum_j B_j$$

Initial values for the attraction factors Bj are needed (see notes given later about the destination choice probabilities).

5.1.3 The composite utilities summed over choices in the time period layer are given by

$$U_{im} = \ln \sum_t \exp(\theta_{time} U_{imt})$$

This uses the scaling parameter θ_{time} which reflects the ratio of the lambda for macro time period to the lambda for distribution.

5.1.4 The incremental composite utilities summed over the main mode layer are given by

$$U_i = \ln \sum_m \exp(\theta_{\mathrm{mod}e} U_{im})$$

These composite utilities are used to calculate the choice probabilities in the various layers as follows. Where required, base utilities can also be calculated from the same composite utility formulae given above, but using base values for the generalised costs and balancing factors.

Choice Probabilities:

5.1.5

Layer 1, Frequency:

$$T_i = \frac{{}_0T_i \exp(\theta_{freq} U_i)}{\exp(\theta_{freq}\ {}_0U_i)}$$

Note that this calculation makes use of a reference utility value

Layer 2, Main Mode Choice (m):

$$T_{im} = \frac{T_i \exp(\theta_{\text{mode}} U_{im})}{\sum_r \exp(\theta_{\text{mode}} U_{ir})}$$

Layer 3, Macro Time Period Choice (t):

$$T_{imt} = \frac{T_{im} \exp(\theta_{time} U_{imt})}{\sum_s \exp(\theta_{time} U_{ims})}$$

Layer 4, Destination Choice (j):

$$T_{ijmt} = \frac{T_{imt} B_j \exp(-\lambda_{dist} G_{ijmt})}{\sum_k B_k \exp(-\lambda_{dist} G_{ikmt})}$$

Notes

All distribution models satisfy the constraint: $\sum_j T_{ijmt} = T_{imt}$

For doubly constrained destination choice models B_j needs to be calculated to satisfy the additional constraint: $\sum_{imt} T_{ijmt} = T_j$

Some models employ area specific, mode specific, and time period specific constants and/or sensitivity parameters which vary by zone or zone pairs. Advice on these matters can be found in sections 1.7.14 to 1.7.16.