Department for Transport

SUPPLEMENTARY GUIDANCE

Bespoke Mode Choice Models

January 2014

Department for Transport

Transport Analysis Guidance (TAG)

https://www.gov.uk/transport-analysis-guidance-tag

This TAG Unit is part of the section **SUPPLEMENTARY GUIDANCE**

Technical queries and comments should be referred to:

Transport Appraisal and Strategic Modelling (TASM) Division
Department for Transport
Zone 2/25 Great Minster House
33 Horseferry Road
London
SW1P 4DR
tasm@dft.gov.uk

Contents

1 Bespoke Mode Choice Models

1.1 Introduction

1.1.1 This supplementary guidance unit discusses the development of bespoke mode choice models as part of the transport demand model. TAG Unit M2 – Variable Demand Modelling discusses the majority of cases where mode choice models are transferred for use in new models for different applications. This unit gives further detail on those cases where it is judged that the development of a bespoke mode choice model for the study is the most appropriate approach.

1.1.2 Mode choice models are used to provide forecasts of passenger demand for transport services. They are typically applied where transport services are new, such as the introduction of a light rail system.

1.1.3 Development of a mode choice model is a specialist activity, requiring skilled and experienced analysts working within structured framework defined by good practice. The objective of this Unit is to provide guidance on the procedures, testing and documentation that would be required for an analyst to demonstrate adherence to good practice. Whilst identifying many elements of good practice, this guidance should not in itself be considered a detailed technical guide to the development of mode choice models.

1.1.4 Bespoke models are more costly to develop than transferred models, because of the data that is required to estimate statistically reliable behavioural parameters. The time required to estimate these models will also be more extensive than that required for model transfer, leading to higher model development costs.

1.1.5 Bespoke modelling will be necessary in the following situations:

- for appraisal of new modes or new characteristics;
- for appraisal of schemes in areas where traveller behaviour may be substantially different from national norms or values from other existing models; and
- the transfer of an entire model system from another area could be considered, only if model transfer is appropriate (see TAG Unit M2).

1.1.6 It is unlikely that anything other than bespoke models should be developed for appraisal of:

- very large public transport schemes;
- schemes which require detailed socio-economic segmentation for appraisal;
- areas where traveller behaviour is substantially different from national norms.

1.1.7 This Unit provides guidance on the procedures and documentation required in the development of bespoke mode choice models. The material covered:

- explains the purpose of bespoke mode choice models;
- provides guidance on model design;
- reviews the stages to data collection;
- gives guidance on model estimation, application and validation; and
- sets out the required documentation for audit.

2 Model Design

2.1 High-level considerations

2.1.1 An important strategic question should be one of 'what is required from the model?' Indeed 'fitness for purpose' should be a guiding principle throughout the design process. Whatever the scheme or policy under investigation, the analyst should always ensure that the budget designated for data collection and mode choice modelling, and the timescale for such activities, is commensurate with the nature and complexity of the problem, and the likely scale and extent of impacts (including any potential second order impacts).

2.1.2 It is important to establish what existing sources of revealed preference (RP), or even stated preference (SP) data, might usefully enhance model development. This includes not only disaggregate choice data but also aggregate planning data and traffic count data. The availability of such data may impact on the scale and scope of any new data collection, the specification of the model (for example, whether the model should accommodate a range of data sources, the extent of segmentation, and the nature of any validation procedure), and the reliability of the model in application.

2.2 Lower-level considerations

2.2.1 Having considered the high-level issues, preliminary model specification should move on to address a series of lower-level issues. Each of these issues is fundamental to model specification and, it follows, data requirements. These issues could have a significant impact on resource needs in analysis.

2.2.2 What is the relevant unit of the decision-maker? For mode choice, this will usually be the individual, although the travelling party may be relevant in some cases. The advantage of the latter is that car cost is accounted for correctly.

2.2.3 What is the choice set? The analyst should identify the alternatives of interest, including any new modes, and take an initial view on the availability of the complete choice set to decision-makers, as well as the extent of any captivity to alternatives.

2.2.4 What are the key behavioural variables of interest? As well as standard variables such as time and cost, the analyst should consider the relevance of 'softer' variables such as crowding/congestion, quality and reliability. This will be dictated by the nature of the scheme or policy under investigation, and the likely scale and scope of its impacts.

2.2.5 What are the key policy variables of interest? These might include fares, service frequency, reliability, accessibility, and quality in public transport modes, and road pricing and parking for cars.

2.2.6 What are the key socio-economic variables of interest? These will usually include age, sex, employment status, income, and car ownership.

3 Model Development

3.1 Logit model

3.1.1 Mode choice models, as conventionally specified, are based on the behavioural principle that a decision-maker will choose the travel mode that yields greatest satisfaction or 'utility'.

3.1.2 Utility is postulated to be a function of both observable (or deterministic) utility and unobservable (or random) utility. Specifically:

$$U_{ni} = V_{ni} + \varepsilon_{ni}$$

where V_{ni} is the deterministic utility derived from alternative i by decision-maker n, and ε_{ni} is the associated random utility.

3.1.3 For purposes of implementation, a specific model form should be adopted. Logit offers substantial versatility; indeed it will be sufficient for many needs. Where more complex forms are deemed necessary, logit offers a valuable benchmark for comparison.

3.1.4 Logit relates probability of choosing alternative i from J alternatives as follows:

$$P_{ni} = \frac{e^{\mu V_{ni}}}{\sum_{nj \in J} e^{\mu V_{nj}}} \qquad (3.1)$$

where μ is a strictly positive scale parameter.

3.1.5 In the context of mode choice, convention is to reinterpret utility as 'generalised cost', which is essentially the negative of deterministic utility expressed in monetary units. The methods discussed below are based on the construct of utility, since this typically includes additional variables - such as ones relating to the decision-maker - that may be difficult to translate into generalised cost terms.

3.1.6 With reference to equation (3.1), the scale parameter μ is inversely related to the variance of random utility (or 'error') as follows:

$$\operatorname{var}(\varepsilon_{ni}) = \pi^2 / 6\mu^2$$

3.1.7 The amount of error has important implications for the properties of the model. All else equal, the greater the error, the smaller the scale parameter and the closer the choice probabilities will tend to $1/J$ for all J alternatives. This issue is known as the 'scale factor problem' and is of particular relevance when estimating models to SP data, which may contain biases and errors typically not found in RP data. Since it cannot in practice be estimated separately from V_{ni}, μ is commonly taken to be one.

3.1.8 It is differences in deterministic utility across alternatives that influence probability - not absolute utility. The relationship of utility difference to logit probability is sigmoid (Figure 1). Thus, if an alternative has an extreme probability (whether high or low), a small change in utility difference will have little impact on probability of choice, whereas if an alternative has a probability close to 0.5, the same change in utility difference will have considerably greater impact on probability.

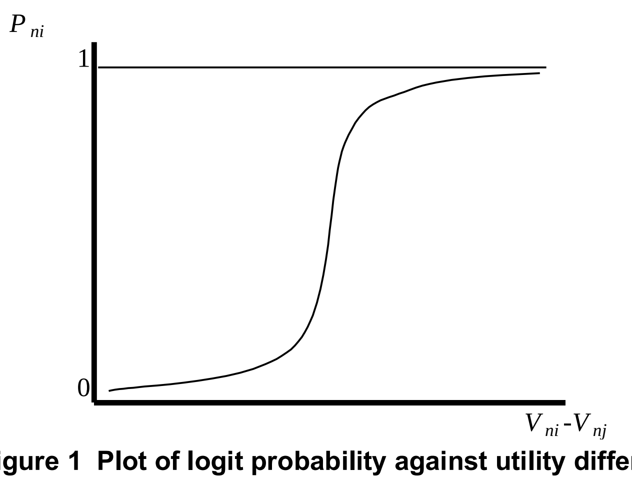

Figure 1 Plot of logit probability against utility difference

3.2 Specifying the utility function

3.2.1 An important practical issue is the specification of V_{ni} . This is typically represented as a function of observed variables relating to the alternative and the decision-maker.

3.2.2 As regards functional form, linear-in-parameters[1] is sufficient for most needs. Indeed, under fairly general conditions, any function can be approximated arbitrarily closely by the linear-in-parameters form.

3.2.3 Variables relating to alternatives may be entered in the function 'directly', as follows:

$$V_{ni} = \sum_k \beta_k x_{nik}$$

where x_{nik} are observations relating to the k th variable (or 'attribute') of decision-maker n and alternative i , and the β are associated parameters. For example, if there were interest in the effect of time (T) and cost (C) on choice, an appropriate representation would be as follows:

$$V_{ni} = \alpha T_{ni} + \beta C_{ni} \qquad (3.2)$$

where α and β are utility or 'taste' parameters relating to time and cost, respectively. Since both time and cost are, in terms of utility, perceived as 'bad', α, β , < 0.

3.2.4 The parameters in equation (3.2) are shown to be 'generic' across choice alternatives. Thus attributes that are common to alternatives are specified as having common parameters, such that estimates of these parameters will be averages across the data. This is not however a requirement; whilst convention is to specify cost as generic, other parameters may be specified not only as mode-specific, but also as person-specific, depending on the focus of interest. Relaxing the assumption of generic parameters allows for different values of time for different modes, people, or both modes and people, for example.

3.2.5 Analysts should adopt the standard of expressing **time in minutes and cost in pence**. Furthermore, care should be taken to express each in terms of single trip units.

3.3 Introducing socio-economic variables

3.3.1 Following from 3.1.7, variables for a given observation that are common across alternatives, such as those relating to the decision-maker, should not be entered 'directly' since they have no impact on utility difference (and therefore probability). They must instead be 'interacted' with variable(s) that do vary across alternatives. For example, if there were interest in the influence of age on the responsiveness of choice to cost, (3.2) could be re-written:

$$V_{ni} = \alpha T_{ni} + \beta_n C_{ni} \qquad (3.3)$$

where:

$$\beta_n = \gamma M_n$$

and γ is the parameter relating to the interaction between cost and age (M).

3.3.2 Substituting into (3.3), the interaction of age with cost can be seen clearly:

[1] Linear-in-parameters means that the model parameters are not cross-correlated between variables, which would be, e.g.: parameter a affects the strength of variables x and y and parameter b affects the strength of variables y and z.

$$V_{ni} = \alpha T_{ni} + \gamma M_n C_{ni} \qquad (3.4)$$

3.3.3 It should be noted that the parameters of (3.4) are again represented generically. Since logit models are usually estimated on data from a sample of decision-makers, it may be revealing to extend the model to investigate the potential for tastes to vary across decision-makers.

3.3.4 If, for example, there were interest in the distribution of tastes with respect to the interaction between age and cost, (3.4) could be re-specified:

$$V_{ni} = \alpha T_{ni} + \gamma M_n C_{ni},$$

which would yield a separate γ parameter for each decision-maker n.

3.3.5 An alternative, and more efficient, representation would be to segment decision-makers by age group, and represent each group by a dummy variable. For any such variable, dummies should be included explicitly in the model for all but one group, thereby avoiding the 'dummy variable trap'. If, for example, the data were assigned to one of three age groups, dummies for two of the groups should be included in the model, with the third acting as the 'base'. The V function should then be represented as follows, where in this case $l = 1, 2$.

$$V_{ni} = \alpha T_{ni} + \left(\gamma_0 + \sum_l \gamma_l D_{ln} \right) C_{ni}$$

and γ_l is the cost parameter specific to segment l.

3.4 Alternative-specific constants

3.4.1 It is essential to include a constant in the utility function of all but one choice alternative. This constant is referred to as to the 'alternative-specific constant' (ASC) or 'mode-specific constant', specifically:

$$V_{ni} = ASC_i + \alpha T_{ni} + \beta_n C_{ni}$$

where ASC_i is the alternative-specific constant relating to alternative i.

3.4.2 The ASC is omitted for one alternative - which becomes the 'base' - again to avoid the 'dummy variable trap'. An ASC can be interpreted as representing the net average effect of omitted variables (relative to the base). The inclusion of ASCs ensures that, when estimated by maximum likelihood, logit is able to replicate the aggregate choice shares.

3.5 Functional form

3.5.1 Thus far, the model has adhered to linearity-in-variables as well as linearity-in-parameters. In some cases, non-linear forms may offer additional flexibility. One such case, which may be useful in mode choice studies, would be to incorporate an explicit 'income effect' i.e. the marginal utility of income diminishes with increasing income. Chapter 8 of Ortúzar and Willumsen (2001) provides direction on such specifications.

3.5.2 As will become evident later, the linear form is particularly attractive when it comes to interpretation of the model, although additional flexibility can be achieved by introducing non-linearity. Among the more popular alternatives are the following:

Quadratic: $\beta C_{ni} + \beta_2 C_{ni}^2$

Power: βC_{ni}^{ς}

Log: $\beta \ln C_{ni}$

Box Cox: $V_{ni} = \beta \frac{\left(C_{ni}^{\lambda} - 1\right)}{\lambda}$, which converges to $\beta \ln(C_{ni})$ as $\lambda \to 0$ and to $\beta\left(C_{ni} - 1\right)$ when $\lambda = 1$

3.5.3 While it is difficult to offer clear prescription, the case for using such forms should be based on a combination of theory (i.e. behavioural rationale) and/or data (i.e. empirical support). One of the more common contexts for non-linear forms is where national-level data (e.g. for trip lengths) may be distinct from local level data.

3.6 Defining the choice set

3.6.1 An important specification task is to define the choice set appropriately. For bespoke mode choice models this will usually be relatively small and clearly defined, at least in an aggregate sense. What may be less obvious is the propensity for decision-makers to consider only a subset of alternatives when actually choosing. There could be any number of reasons why particular alternatives might not be considered, although by far the most common issue in mode choice modelling is car availability. It is important to identify such constraints, and represent the appropriate choice set for each decision-maker (an ‘unavailability of alternatives’ command is provided in most software packages). Although it is necessary, for successful estimation, that at least some decision-makers choose each choice alternative, it is **not** a requirement that all decision-makers have access to the full choice set.

3.6.2 Where choice models are developed to consider the demand implications of a new mode, there are a range of issues associated with data collection and how the new mode should be considered during forecasting. As regards the former, a requirement for a SP experiment would often be implied (guidance is offered in section 2). As regards the latter, a particular issue is the specification of ASCs (Sections 3.4 and 5.3).

3.7 Independence from Irrelevant Alternatives (IIA)

3.7.1 In defining the choice set, it should be noted that logit is characterised by the property of independence from irrelevant alternatives (IIA); that is, for any two alternatives, the ratio of their choice probabilities is unaffected by the presence or absence of any other alternatives in the choice set. Where two alternatives in the choice set are closely related in some sense (i.e. the ‘red bus-blue bus’ problem), IIA is violated, and the use of logit is (in principle) inappropriate. It should be noted that the IIA property of the logit model is evident at the level of the decision-maker and not always present for groups of decision-makers.

3.7.2 There are a number of ways to identify cases where the IIA assumption is violated, but arguably the most practical is to calibrate nested models. This process not only identifies whether IIA applies, but also how to alleviate it (see Section 3.12).

3.7.3 Where IIA provides an accurate representation of reality, it may permit considerable efficiency in analysis, since models can be estimated on restricted choice sets.

3.8 Maximum Likelihood estimation

3.8.1 Logit can be estimated on RP data, SP data, or a combination of the two. Such data are usually collected on a sample of decision-makers from the population of interest. Data collection is considered in detail in section 4.

3.8.2 Convention is to estimate logit by maximum likelihood (ML), the purpose of which is to estimate the parameters for which the observed sample is most likely to have occurred. A number of software packages offer routines for ML estimation, although these may vary considerably according to their cost, ease-of-use and flexibility. Whichever software is chosen, estimation of logit by ML is usually reliable, and it is uncommon for close examination of the ML routine to be required.

3.8.3 Where computational problems are encountered in estimation, closer examination of the ML routine may be necessary. A reasonably detailed account of the most popular ML algorithms is offered in Train (2003), along with diagnostic advice on how common estimation problems may be overcome.

3.8.4 Having estimated a logit model by ML, an initial post-estimation check is needed to ensure that the ML routine converged successfully. If the model failed to converge, then it is necessary to investigate the reasons for this, resolve them, and repeat the estimation. In the event of such problems, the software may provide appropriate prescriptive advice, although it is often necessary for the analyst to interpret such advice in the context of how the data, model and estimation routine have been specified. **The analyst should not draw behavioural conclusions from software failure.**

3.9 Preliminary interpretation

3.9.1 Having estimated logit successfully, a series of preliminary tasks in statistical inference should be undertaken. Each of the utility parameters should be subjected to a Student's t-test for statistical significance and, strictly speaking, only parameters that are of statistical significance should be retained for purposes of model application. Most software packages capable of multinomial logit estimation will do this as a matter of course. In practice, however, the decision to include/exclude a given variable is less clear cut and depends as much on the sign and relative magnitude of the coefficient as well as its standard error. Accepting a coefficient with an inappropriate sign or magnitude simply because it is statistically significant is clearly wrong, as is rejecting a key policy variable if it is marginally insignificant. The development of a choice model is largely guided by experience, informed by what the standard errors infer about the accuracy of the coefficients.

3.9.2 When making such judgements, it may be informative to consider the relationship between statistical significance and sample size. More specifically, for large populations and relatively small samples - which is the typical context for mode choice modelling - the standard error of an estimate relates approximately to sample variance and sample size as follows:

$$se\left(\hat{\beta}\right) \approx S(\beta) \big/ \sqrt{N}$$

where $S(\beta)$ is the sample variance of β and N is the sample size. To illustrate this relation, a quadrupling of the sample size would, for given sample variance, imply a doubling of the t-ratio in a test of statistical significance.

3.9.3 Such considerations may also impact on the specified degree of segmentation, since greater segmentation may imply reduced standard errors for segment-specific parameters. Moreover, where budget and/or other constraints restrict sampling, the retention of insignificant variables may be justifiable if a modest expansion of the data set is likely to bring significance.

3.9.4 The sign of each significant parameter should be assessed as to its intuitive validity; for example, fares should always have a negative effect on utility.

3.10 Further interpretation and diagnostic testing

3.10.1 Analogous to least squares estimation, the prevalence of any (near) colinearity between variables may affect the sign and/or significance of parameter estimates. Such dependency can be investigated through estimating models with restricted sets of variables, and examining the behaviour of the model as variables are added or removed. Good estimation software will also produce parameter correlation matrices, analysis of which will inform any such investigations.

3.10.2 Referring back to 3.1.6, it should be noted that the parameters in the utility function are scaled relative to the variance of unobserved factors; larger variance in ε_{ni} will lead to smaller . When it comes to interpretation, therefore, ratios of parameters are more meaningful than β absolutes, since the scale factor μ cancels out.

3.10.3 A further attraction of ratios of parameters is that, at least in the case of a linear functional form, they have ready economic meaning as 'marginal rates of substitution'. In particular, if the denominator of such a ratio is a cost parameter, then the ratio can be interpreted as the marginal rate of substitution with respect to cost, or in other words 'value'. For example, and with reference to (3.3), the value of time is given by the ratio of time and cost parameters:

$$VOT = \alpha / \beta$$

3.10.4 VOT can, more generally, be derived from any functional form by taking the ratio of marginal utilities, as follows:

$$VOT = \frac{\partial V / \partial T}{\partial V / \partial C}$$

3.10.5 Any derived valuations should be tested for statistical significance; tests for significant difference from 'reference' values (such as 'standard' values') may also be insightful.

3.10.6 If estimated by ML, the goodness of fit of a logit specification should be measured using the log-likelihood (commonly referred to as 'rho-squared') index. The basic form of this index is defined:

$$\rho^2 = 1 - LL_f / LL_r$$

where LL_f is the final log-likelihood of the full model, and LL_r is the final log-likelihood of a restricted model.

3.10.7 Although a number of restricted models may offer bases for meaningful tests, a minimum requirement should be to implement the test with a market share model (i.e. a restricted version of the full model that includes only ASCs) as the base. Such a formulation yields the widely used 'rho-squared with respect to constants' index.

3.10.8 The ρ^2 index offers a measure of the goodness of fit of the logit model, and is analogous to the R^2 statistic in ordinary least squares regression. The value ρ^2 lies between zero and one, but values between 0.2 and 0.4 are often considered indicative of very good fits. In common with R^2, the ρ^2 with respect to constants is comparable across different samples.

3.10.9 Expanding on the t-tests for hypotheses regarding individual parameters, it may be insightful in some cases to test more complex hypotheses regarding subsets of parameters. Two of the more common such hypotheses are (i) that the coefficients of a subset of variables are collectively zero; (ii) that the coefficients of two variables are the same. Both of these tests can be implemented using a likelihood ratio test, which is given by the general form:

$$R = L_r / L_f$$

where L_r is the final likelihood of the restricted model under the null hypothesis (e.g. in case I, the restricted model would constrain the relevant subset of coefficients to be zero), and L_f is the final likelihood of the unrestricted model. Thus a restricted model should be estimated by ML in accordance with the null hypothesis. The test statistic is given by $-2\log R$, which is distributed chi-squared with degrees of freedom equal to the number of restrictions implied by the null hypothesis.

3.11 Validation

3.11.1 Having conducted the above procedures in statistical inference, the properties of the estimated model should be validated against benchmark empirical evidence. Such investigations should focus on two principal constructs - valuation and elasticity. As regards the former, any valuations implied by the estimated model should be reconciled, where possible, with empirical evidence from comparable local schemes. Where such evidence is unavailable at the local level, recourse to national evidence should be made. As regards the latter, the elasticity properties of the model should be similarly compared against available local evidence. Such analysis can be based on measures of point elasticity, calculated for both direct and cross effects, across the sample. The relevant formulae for direct and cross elasticity for each decision-maker are, respectively:

$$E_{ix_{ni}} = \frac{\partial V_{ni}}{\partial x_{ni}} x_{ni} \left(1 - P_{ni}\right)$$

$$E_{ix_{nj}} = \frac{\partial V_{nj}}{\partial x_{nj}} x_{nj} P_{nj}$$

3.11.2 To obtain elasticity estimates for the sample as a whole it is usual to take a weighted average across the elasticity estimates for each decision-maker, with the weights being the individual choice probabilities for the mode in question. Simply inserting average values for P and x will, if there is any variance to the data, lead to an aggregation bias and incorrect elasticities. An alternative method is to make small changes to the variables during model application, and derive arc elasticity estimates from the predicted market shares.

3.11.3 The validity of the estimated model should be further tested in implementation. The estimated model should be applied to forecasting (the subject of forecasting is considered in section 5), and the ability of the model to replicate observed market shares assessed. A range of indicators of forecasting performance may be employed, although a simple and robust test is offered by a Chi-squared test:

$$\chi^2 = \sum_{j=1}^{J} \frac{\left(f_a - f_p\right)^2}{f_p}$$

where

f_a is the actual frequency

f_p is the forecast frequency

J is the number of alternatives in the choice set

with degrees of freedom:

$$df = J - m - 1$$

where m is the number of parameters to be estimated on the basis of the sample data.

3.11.4 The validation process should be carried out across a number of dimensions including those defined by the characteristics of the sample (e.g. income, gender, age) and the attributes of the choice alternative (e.g. cost, in-vehicle time).

3.12 Nested logit

3.12.1 With reference to Section 3.7, a diagnostic for, and (partial) resolution to, the property of IIA is offered by the nested logit model, which groups similar alternatives together in mutually exclusive subsets or ‘nests’ (i.e. an alternative can be included in only one nest). Choice probability is represented as the product of marginal probabilities of choosing nests and the conditional probability of choosing a given alternative from a nest.

3.12.2 Nested logit can be illustrated by considering a problem of two-levels and two-nests, although the model can in principle be extended to any numbers of levels and nests. With reference to the tree diagram in Figure 2, choice probability is given by:

$$P_i = P_m P_{i|m}$$

where P_m is the marginal probability of choosing nest m, and $P_{i|m}$ is the conditional probability of choosing alternative i from nest m.

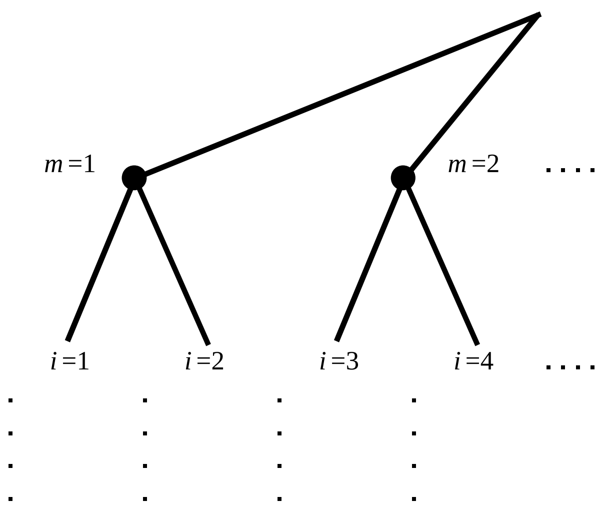

Figure 2 Nested logit for a two-level two-nest problem

3.12.3 Within each nest, the property of IIA holds, and conditional probability is represented as logit:

$$P_i = \frac{e^{\frac{\mu V_i}{\mu_m}}}{\sum_{i' \in J} e^{\frac{\mu V_{i'}}{\mu_m}}} \qquad (3.5)$$

where μ_m is the scale parameter relating to nest m.

3.12.4 Turning now to marginal probability, the utilities from (3.5) are introduced in an expression for the Expected Maximum Utility of each nest m (commonly referred to as the ‘log sum’ or composite cost), as follows:

$$I_m = ln \sum_{i' \in m} e^{\frac{\mu V_{i'}}{\mu_m}}$$

3.12.5 The probability of choosing an alternative in nest m is also of the logit form and is shown as:

$$P_m = \frac{e^{\mu_m I_m}}{\sum_{m' \in M} e^{\mu_m I'_m}}$$

3.12.6 A common simplification is to assume that μ_m is constant across all m in a given level of the tree (i.e. all nests at a given level have the same scale factor).

3.13 Estimating and interpreting nested logit

3.13.1 It is very important to note that there are (in general) two different and commonly used specifications of the nested logit model. In particular some applications are specified and estimated without dividing the lower level utility by μ_m .

3.13.2 Since this distinction between specifications may have substantive implications for interpretation and application, the analyst is advised to seek appropriate advice from the software supplier before proceeding.

3.13.3 The inferential and diagnostic analysis required following estimation is essentially the same as for logit, although one additional test is required in order to check the internal consistency of the nested logit structure. Following from 3.13.2, the precise specification of this test differs according to the specification of nested logit adopted. For example, if lower-level utility is specified without dividing through by μ_m , the test requires that:

$0 < \mu_m \leq 1$ for all m (3.6)

3.13.4 Interpreting (3.6), where μ_m is not significantly different from one, there is a violation of the IIA assumption (see Section 3.7); where this holds for all μ_m the nested logit model collapses to logit.

3.13.5 Although the above discussion was based on a two-level problem, the analysis can be readily extended to more than two-levels, with different scale factors at each level. The internal consistency test then involves an extension of (3.6).

3.13.6 A practical difficulty with nested logit is that the most appropriate nesting structure may not always be obvious. It may therefore take some effort to identify a definitive structure, judgements on which should be based on internal consistency, relative explanatory power and other properties of the model such as implied valuation and elasticity.

4 Data Collection

4.1 Available Data

4.1.1 Revealed Preference (RP) refers to observations of actual behaviour, for example the mode choices that decision-makers currently make or made in the past. RP data is inherently more credible than SP data and its use, if only partially, will strengthen the credibility of demand forecasts in the appraisal framework. There are three types of Revealed Preference (RP) data that could be used for mode choice modelling:

- aggregate data: including information on aggregate mode shares, mode shares by trip length, etc.;

- semi-aggregate data: reflecting proportions of choices made by groups of travellers, typically from matrix data of choices by origin, destination and purpose categories, trip length distributions, etc.;
- disaggregate data: reflecting detailed observations of actual mode choice behaviour, for a sample of travellers, in the relevant study area.

4.1.2 In addition count data can provide information on aggregate mode shares.

4.1.3 For all mode choice models, aggregate RP data are required as a minimum, either for recalibration or validation of mode shares. Information on mode shares by trip length could be used to (manually) calibrate the model scale, but semi-aggregate or disaggregate data provide much better information for this purpose. It is therefore recommended that data on mode shares by trip length category is the minimum information required to calibrate the model scale and is appropriate for development of models for appraisal of small schemes only.

4.1.4 Semi-aggregate data can be used to provide appropriate scale and constant adjustments. For bespoke models, it is recommended that semi-aggregate data be supplemented with additional disaggregate RP or SP data, in order to estimate statistically reliable model parameters.

4.1.5 Disaggregate data is much richer than semi-aggregate data, in terms of its explanatory power for modelling, and is therefore preferable for bespoke modelling (and for calibration of transferred models, although often disaggregate data will not be available, which may be a reason for opting for a model transfer). Disaggregate choice data can be collected from en-route postcard surveys, from home or phone interviews, travel diaries, as well as from existing sources such as the National Travel Survey. Data collected using choice-based survey procedures, for example, interviewing passengers on a bus or train, will lead to samples which are not representative of observed mode shares. Adjustments to take account of these biases must therefore be made in the model estimation procedure.

4.1.6 Stated Preference (SP) refers to observations of hypothetical behaviour under controlled experimental conditions. A scheme that introduces a new mode, for example, would imply a need for SP analysis, since RP data is by definition unavailable for such a context. When developing a bespoke mode choice model therefore often requires new SP analysis. Stated Preference (SP) data can also play an important role in bespoke modelling, particularly for modelling demand for new modes (see Section 0 for a detailed discussion of the role of SP data).

New Revealed Preference (RP) Data

4.1.7 New RP surveys should collect, as a minimum, the respondents' origin, destination, choice of mode and purpose of travel.

4.1.8 For public transport users, information on licence holding and car availability is important and for users of all modes, pass holding or entitlement is valuable. Other background characteristics, for example; age, gender, income and employment status are desirable but may not be possible to collect in the context of the survey.

4.1.9 For mode choice models, observations of people choosing the specified modes are required for all (existing) modes considered in the study.

Revealed Preference and Stated Preference

4.1.10 RP data can be obtained from SP respondents, from postcard surveys, home or phone interviews, travel diaries, as well as from the National Travel Survey and Census.

4.1.11 The collection of RP data is not without problems however. There are often large biases in respondents' self reported data, underestimating the costs of their chosen mode and overestimating the costs of alternative modes. To overcome these problems it is sometimes necessary to use explanatory variables from network models and published timetable data. Even where respondents' reported data is modelled, there is often a considerable amount of missing data which needs to be collated.

5 Model Application

5.1 Introduction to model application

5.1.1 Logit and nested logit are usually estimated on probabilities of choice for a sample of decision-makers. What is typically of interest to policy-makers, however, is an aggregate measure of these probabilities - i.e. market share - across a population.

5.1.2 The application of average measures of explanatory variables to the calculation of probability yields biased measures of average probability.

5.2 Sample enumeration

5.2.1 Consistent estimates of market share can be obtained using sample enumeration. This involves calculating, for each decision-maker in a sample, the probability of choice for each alternative in the choice set. These probabilities are then aggregated over decision-makers; average probability can be obtained by dividing through by the sample size.

5.2.2 More formally, a consistent estimate of the number of decision-makers choosing alternative is given by:

$$\hat{N}_i = \sum_n w_n P_{ni}$$

where w_n is the weight attributed to decision-maker n. The w_n parameter represents the number of decision-makers similar to decision-maker n in the population, i.e. the number of decision-makers within each segment of interest. Thus if the sample is random then w_n is constant for all n, whereas if the sample is segmented then w_n is the same for all n within a segment. If the sample is not representative of the population, then the weights should be adjusted accordingly.

5.3 Adjusting the ASCs

5.3.1 In applying a model with ASCs to forecasting, it should be recognised that the influence of explanatory variables not represented explicitly in the model may change between estimation and forecast contexts (e.g. over time). Such changes can be accommodated through re-calibration of the ASCs. This involves inserting the estimated parameters (including the ASCs) in the model, along with the base data, and assessing the ability of the model to replicate 'target' market shares. If the forecast shares differ significantly from the target shares, then the ASCs should be adjusted, and the analysis repeated iteratively.

5.3.2 Target market shares may be based on external evidence or by particular requirements relating to a forecast segment. In the latter case, for example, there may be an interest in the ability of the model to forecast accurately for a particular segment of the sample, and a need to tailor the ASCs accordingly. Adjusting the model constants for existing modes is relatively straightforward as the base market shares will be known. Setting the ASC for a new mode is however more problematic, as the values from SP research will be estimated to choice sets different from those to which they are applied, may be of the wrong scale, and are likely to be subject to various respondent biases inherent in the SP survey There is no easy solution here, and recourse to similar travel situations may be required. The constant for the new mode is therefore a strong candidate for sensitivity testing.

5.4 Forecasting

5.4.1 Forecasting involves applying the above aggregation methods to some alternative scenario, defined on the basis of two inputs:

- data on the utility variables under the scenario of interest (e.g. reflecting an increase in fares); and
- second, the w_n parameters (e.g. reflecting changes in socio-demographics).

Changes to the latter are particularly important for long term forecasts, where changing patterns in population, income and car ownership are likely to be influential on demand.

5.5 Patronage build-up

5.5.1 In most instances, the mode choice model will predict an equilibrium state in which mode switching occurs instantaneously (e.g. in SP). In reality, however, there is likely to be inertia within the market, perhaps because of dissipation of knowledge about the service and/or a delayed behavioural response to the new journey opportunities (e.g. in RP). A prudent forecaster might factor down initial patronage forecasts to take account of the delay in take-up. This can be done 'off-model' using rules of thumb or included within the model by means of an inertia term that decays over time. In the long run (greater than 2 years), one would expect the overwhelming degree of inertia to have disappeared.

6 Reporting

6.1.1 A thorough audit trail should be completed during model development to justify the methodological approach taken and any assumptions that are made. Guidance for the Technical Project Manager describes the reporting requirements expected for a major scheme business case. Reporting of the development and operation of the demand model should e described in these reports, particularly the **demand model report**.

6.2 Data collection Report

6.2.1 The **data collection** report should include the following:

- Documentation of the key findings of focus groups (if undertaken);
- The RP data collection exercise and sampling strategy;
- The SP data collection exercise and sampling strategy. This will also contain information on the questionnaire design, testing by simulation and pilot survey results; and
- Data processing and cleaning. This will include information on the processing of raw data for use in final model development.

6.3 Demand Model Report

Model Design

6.3.1 The demand model report should include the following considerations of model design:

- Information on the nature of problem and the objectives of the likely solutions;
- A definition and size and scope of the study area;
- The availability of existing data to establish new models;

- The need to undertake new surveys to establish new models;
- Preliminary model specification, including information on model structure, explanatory variables and estimation procedures;
- Details of the software to create and apply the model;
- The forecasting parameters and years for which the forecasts are required; and
- Information on the timescale and resources required for model development.

Model Estimation

6.3.2 The demand model report should include information on model estimation:

- the specification of logit models calibrated to each data set;
- the specification and justification for alternative nested structures; and
- the merging of different data sets to develop joint RP-SP choice models.

For each model, evidence and justification is required for:

- the inclusion/exclusion of each variable;
- the specification of the functional form;
- the degree of market segmentation; and
- the significance of any structural coefficients.

For each reported model, information should be presented on:

- the variables included, their unit of measurement, and which alternatives they apply to;
- the estimated coefficients and associated t-statistics/standard errors. Where models are estimated to SP data, the standard errors should be adjusted to account for repeat observations;
- the number of observations; and
- the explanatory fit of the model;

And where appropriate:

- the relative attribute valuations (e.g. value of time) together with estimates of their statistical confidence; and
- the implied elasticities of demand.

Model Application

6.3.3 The demand model report should include information on model application:

The models should be applied where possible using sample enumeration techniques. Documentation is required to:

- Justify the approach to model application;
- Report any weighting of the sample to make it representative;

- Show how explanatory variables change over time; and
- Show how the model is able to recover the base market demand forecasts over a range of market segments

Model Validation

6.3.4 The demand model report should include information on model validation:

- The report should comment on the credibility of the forecasts when compared to actual patronage figures for similar schemes;
- The report should review the relative attribute values (e.g. value of time) implied by the model and compare them to published evidence; and
- The report should review the own and cross elasticities of demand implied by the model and compare them to published evidence.

6.4 Forecasting Report

6.4.1 The forecasting report should include

- Forecasts of generalised cost, passenger demand, revenue and kilometrage by O-D pair and mode;
- Estimates of patronage build-up over time;
- Sensitivity tests on key input parameters; and
- Specification of the schemes tested and scenario forecasts

7 References

Bateman, I. et al. (2002) Economic Valuation With Stated Preference Techniques: A Manual. Edward Elgar, Cheltenham.

Kocur, G., Adler, T., Hyman, W. and Audet, E. (1982) Guide To Forecasting Travel Demand With Direct Utility Measurement. UMTA, USA Department of Transportation, Washington D.C.

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000) Stated Choice Methods: Analysis and Application. Cambridge University Press, Cambridge.

Ortúzar, J. de D. and Willumsen, L. (2001) Modelling Transport, 3rd Edition. John Wiley, Chichester.

Ortúzar, J. de D. (2000) Stated Preference Modelling Techniques. PTRC, London.

Train, K. (2003) Discrete Choice Methods With Simulation. Cambridge University Press, Cambridge.

Pearmain, D. and Kroes, E. (1990) Stated Preference Techniques: A Guide to Practice. Steer, Davies & Gleave Ltd., Richmond.

Passenger Demand Forecasting Handbook (ATOC)

TRL Report TRL593, The demand for public transport: a practical guide. http://www.demandforpublictransport.co.uk/TRL593.pdf

8 Document Provenance

This unit is based on former TAG Unit 3.11.4, which itself was based on Major Scheme Appraisal in Local Transport Plans Part 3: Detailed Guidance on Forecasting Models for Major Public Transport Schemes.