Task 7.2 Trip Production Statistical Analysis

Important note
There is a strong mover towards trip end modelling on a person rather than a household basis in international modelling practice. In any modelling approach some link between planning data (generally person-based) and the trip end model is required, referred to herein as the family structure model); this is a significant complication (see Task 7.6).

While the descriptions below appear to make the person model seem complicated, in fact the model itself is extremely simple, comprising a trip rate for each person, the rate varying by person characteristics. The rate may also vary by household characteristics, although this seems only likely to be an issue for a few trip purposes.

Inputs
Estimation files.

Processing
Choose statistical software: eg SYSTAT or LIMDEP.

The model specifications below are non-standard but draw on other model specifications and the generally consistent findings on influential variables to create a convenient forecasting framework.

Although the current model makes use of household trip rates, we strongly recommend a movement over to **person** trip rates, while retaining the facility to include household effects where justified.

Home-Based Work (HBW) Model

Variations in trip rates are likely to be as follows:

- person effects: a function of whether or not the person is employed; also expect the trip rate to vary by work structure (full/part-time; contractor ...[1]) because these dictate the need to make a commuting journey;
- household effects: number of cars possible but unlikely; inverse correlation with number of children possible (the school trip substituting for the work trip), but this seems too detailed a refinement and of little policy interest; possible correlation with location, but again unlikely.

Conclusion: we should seek a person trip rate model sensitive primarily to work structure.

Model form:

- proposed model relates work trips to full/part time, work arrangements and type of employment

$$\text{Work trips/zone} = \Sigma_{at} [NF_{at}*WF*\alpha_{at} + NP_{at}*WP*\beta_{at}]$$

where:

NF_{at} NP_{at} are the number of full and part time employed residents in the zone in work arrangement category 'a' and type of employment category 't', and potentially by car ownership segment

WF WP are mean person trip rates for full and part time employed persons respectively, and potentially by car ownership segment

[1] In some models, while and blue collar workers are distinguished, but we are not convinced of the usefulness of this segmentation.

α_{at} β_{at} are trip rate adjustment factor matrices accounting for the effects of work arrangements and type of employment, the optimum number of such factors to be determined

- estimation simply involves determining the mean trip rates and factor matrices; for the latter, we need to determine which of 16 combinations of 'a' and 't' are significant and would expect to compress the factors from the maximum of 16 to a very few significant effects;
- we may test whether WF & WP are functions of other household or person characteristics although this would significantly complicate the forecasting and there is little evidence that they would be significant.

Whether or not the trip rate variations by car ownership are significant, we need to split these productions into choice/competition/captive, according to the definition established in "Segmentations". This requires the number of work trips to be apportioned between these categories in proportions p_{ca}, p_{co}, p_{ch} (it needs to be determined whether these proportions should vary between full and part time).

Home-Based Education (HBEd) Model

Similar to the HBW model, variations are likely to be primarily due to person type:

- person effects: expect to vary by age of child (essentially starting to reduce from school leaving age); then lower rates for young adults in higher education; then tiny rates for older adults;
- household effects: no particular interactions are expected with household characteristics.

Model form:

- proposed model:

  Education trips/zone = $N_{6\text{-}16}$*$ET_{6\text{-}16}$ + $N_{17\text{-}25}$*$ET_{17\text{-}25}$ + $N_{>25}$ * $ET_{>25}$

  where:

  N_i is the number of persons in age group 'i' in the zone, and

  ET_i is the education trip rate for persons of that age group
- estimation involves determining the mean trip rates and identifying the optimum age classes based on schooling regulations, and/or by tabulating schooling probabilities by age group using the household survey, census or education statistics; issue of pre-school and kindergarten;
- we may need to split these into choice/competition/captive.

Home-Based Shop (HBSh) and Home-Based Social (HBSo) Models

These models should be tested individually and together because of the large overlap in these purposes (when is shopping classified as shopping and when recreation and/or social?). Expected trip rate variations are:

- person effects: expect to be a function of the 7 person categories;
- household effects: expect to be a function of mobility (ie car ownership), household size and/or structure (for example, shop trips may be a household activity and the person trip rates may thus reduce with increasing household size) and/or location (accessibility to 'attractions');
- in an attempt to limit the analysis, we might expect the following relationships:
  - with a small household survey sample, we should not expect or seek substantial categorisation and segmentation of trip rates[2];

[2] Income is an interesting variable with high correlation with car ownership, person type and family size and posing additional forecasting complications.

- person trip rates will mainly vary by person type and household car ownership, and this would be the basic segmentation; we might reasonably hope that the car ownership effect is uniform across person types (ie that the effect of car ownership on mobility is to increase uniformly the trip rate of all persons in the household);
- there may be secondary effects of location (eg urban/rural) and household size (for the reasons given above; 1, 2+ could be the major distinction); other effects might be number of adults and household employment status (unemployed household, pensioner household, employed household), but this seems as though it would be pushing the potential of the data for trip rate segmentation too far.

Model form:

- the basic person trip rate model described above has the following form:

  HBSh/HBSo trips/zone =

  $\Sigma_i N_i * T_i + \Sigma_j N_j * \Delta T_j$

  where:

  N_i are the numbers of persons in the zone for each person type i,

  N_j are the number of persons in the zone in households of car ownership level j = 0, 1, 2, 3+,

  T_i is the average trip rate for person type i,

  ΔT_j is the incremental mobility effect on the person trip rate for households of car ownership level j
- further incremental trip rate effects could be added for household size and location using the same formulation; if interaction effects were observed (between these incremental variables), the model would become significantly more complicated.

Non Home Based Other (NHBO)

This model would appear to have the same form as the HBSh and HBSo models although, there being less travel data, it is unlikely to support as detailed a segmentation or structure.

Business (BU) Model

This model would seem to have an identical form to the HBW model, but it may be worthwhile considering a further segmentation based on occupation and/or industry;

It is separated because the high value of time is important for evaluation and significant for tolling studies (this is worthwhile if business trips form a significant market and can be modelled reasonably reliably).

Basic Calibration Steps

- predefine set of model specifications to be considered (with assistance of tabular analyses);
- use statistical analyses to establish preferred models;
- analyse models for geographical bias/fit based on some predefined geographic aggregations and/or the sector aggregations; incorporate any geographic k-factors[3];
- produce range of model calibration fit statistics (eg R^2, T-statistics, predicted vs observed plots at zonal level and also for sector aggregations);
- report.

[3] The best models of this type do not reproduce many of the geographic variations in the data. It is therefore important that these are identified and checked and factors established to correct for significant differences.

Outputs
Trip production model.
Report.