Task 1.5 Roadside Survey Editing and Processing

The processes described below may be undertaken within a relational data base, many tasks being done simultaneously.

1.5.1   Cleaning and Editing Interviews and Counts

Inputs
Roadside questionnaire data base
Manual Classified Counts (MCC)
Automatic Traffic Counts (ATC)

Processing
For interviews this involves:

- defining MIC

- range checks on questionnaire responses

- if possible, checking on reasonableness of address data (origin, destination and survey site grid references can be used to check that a trip with this origin and destination was likely to have passed through the survey site)

For counts it involves:

- MCC: checking for completeness and miscoding
- ATC: checking for tube failure and missing data

Outputs
Cleaned roadside questionnaire data base
Cleaned MCC & ATC files

1.5.2   Processing and Expansion

Inputs
- Cleaned roadside questionnaire data base
- Cleaned MCC & ATC files

Processing
(i) Set up common purpose codes.

(ii) Deal with MIC issues.

(iii) Survey period expansion factors should be appended to the trip records as follows:

- check interview sample by site, day and time period; determine whether to use 15min, 30min or 1 hour basis for the expansion period (this decision will vary by vehicle type with commercial vehicle expansion being on a more aggregated basis because their numbers are less);

- aggregate MCC counts by site, day and time period;

- tabulate number of clean interviews by site, day and time period;

- aggregate time periods where number of interviews is too few in the basic time periods to reliably expand (eg expansion factors would be too big);

- compute expansion factors as ratio of counts to interview sample for each site, day and time period;

- append expansion factors to interview records;

– verify process by tabulating expanded data and comparing with counts and/or assigning to network.

(iv) Expansion to 24 hours is done by adding to the trip records a 24 hour factor by direction derived from ATC data. At this stage, survey day bias can be corrected by applying a factor based on the survey day 24 hour traffic volume to the average weekday traffic volume in the survey week(s).

(v) If there is any chance of trip double-counting between the survey sites, then this should be eliminated (it applies if some car trips could potentially be intercepted at more than one site) by applying a double-counting factor of 0.5 to all such journeys[1].

(vi) Out of scope[2] data needs to be excluded. For Wellington, we processed the external roadside survey data to eliminate residents' home-based trips travel (which was encompassed by the household survey), identify non-residents' home based trips, and isolate non home based trips. The latter trips could be either by residents or non-residents so a factor was derived to exclude residents' trips (which were covered by the household survey), based on the residents/non-residents split for home based external trips (by time of day).

(vii) The roadside interviews were in both directions of travel so trip reversal was not required. If the survey is carried out in only one direction (see the rail survey processing description) then it will usually be necessary to synthesise from the survey data the characteristics of the reverse flow, and the following procedure may be required:

– the survey records are duplicated, and the O and D zones and purposes are swapped;

– for the interview direction, the time of travel is the time of interview;

– for the return trip, the time of travel is that provided on the questionnaire by the respondent;

– for some questionnaires the return journey time will not be provided, either by omission or because the trip is 1-way; an appropriate process is to be developed to account for this, for example:

  o the reversed interviews for which a time of travel is available can be expanded by hour to the reverse counts;

  o if the irreversible trips have different characteristics to the others then extra bias factors may be applied to the reversed trips, based on the interview-direction time period and trip purpose, to account for the specific characteristics of the trips which cannot be reversed.

(viii) An additional factor to expand the survey from the daylight survey period to the full 24 hours was determined from the ATC data.

Outputs
Cleaned roadside interview data base with expansion factors appended.

---

[1] And triple counting by a factor of 0.33 etc. The factor assumes similar sample variances - if some sites have higher sampling rates than others for particular journeys, then the method of combination is best based on the sample variances (using a weighting based on the inverse of the variance).
[2] This refers to data which is not required because, for example, it is fully external to the study area or because it is being sourced from another survey.