# Task 2.12 Use of Intercept Matrix Data & Combined Data Processing

## Introduction

There are four activities to be specified:

- providing trip data for model estimation, which requires a decision on what data to use for this purpose;
- providing a base road matrix for assigning to obtain costs for model calibration, and in the process checking for bias in the observed trip data;
- providing best estimate matrices for model application.

## Available Travel Survey Data Sources

| Purpose/Mode | Car | Public Transport |
|---|---|---|
| **HBWork** | Household | Household<br>Rail Passenger |
| **HBEducation** | Household<br>School | Household<br>School<br>Rail |
| **Other purposes** | Household | Household<br>Rail Passenger |
| **Resident external trips** | Household<br>External roadside | N/a |
| **Non-resident external trips** | External roadside | N/a |

N/a: not applicable

## Establish General Data Consistency

All of the comparisons of data sources which follow are envisaged to be based on:

- an aggregated TLA to TLA[1] expanded matrix; for this we need some concept of the sampling errors associated with each survey (for this I suggest we use my survey sampling error spreadsheet, assuming uniform sampling for each survey, and assume that the variance of the difference between 2 samples = sum of the variances of the individual samples and use a simple t-test of 95% significance);
- a frequency distribution of matrix cells by number of sampled (unexpanded) trips (to compare survey coverages).

Comparisons:

- public transport trips: compare volumes in a TLA-TLA matrix and trip frequency distributions, for each purpose:
  - comparison of rail trips in household and rail surveys by purpose;
  - HBEd: comparison of education school and public bus trips by school age children between the household and school surveys.

- HBEd car trips: while we could compare education trips by school age children between the household and school surveys, this would be complicated as the household survey also encompasses car passenger and escort trips; as we have no intention of using the car trip data from the school survey in the matrices, we shall omit this comparison.

[1] TLA: territorial local authority, of which there are 8 in the Wellington region.

Conclusions from this analysis will establish/confirm:

- that the intercept data provide much better matrix coverage, and will quantify this;
- whether the surveys are consistent at the TLA level, ie the numbers of trips are not statistically different.

## Best Estimate Public Transport Matrices for Model Estimation and Application

Public transport trips have low samples in the household survey. However, this is not the case for our other sources of public transport trip data, the school and rail surveys (which I will refer to as 'intercept' data). We need a means of improving the household survey matrices on the basis of the higher sample intercept data.

- **Sampled PT trips in three surveys are shown in the table.**

| Survey | Mode | | |
|---|---|---|---|
| | **Rail** | **Public Bus** | **School Bus** |
| Household | 724 | 827 | 338 |
| Rail | 5500 | - | - |
| School | 161 | 259 | 445 |

What follows is on a 'main mode' basis, so avoiding double-counting.

We seek to create a best estimate observed public transport matrix by combining best estimate rail and bus matrices.

The best estimate rail matrix is simply that from the rail survey – there is nothing to be gained from combining this with the household or school surveys whose samples are so much smaller.

For public bus (and school bus if needed) the household and school data can be combined for school age education trips[2] using an inverse variance approach (presuming a fully consistent HBEducation trip definition for the two surveys) then added to the other public transport trips from the household survey.

The bus and rail matrices can then be summed.

For the above we need the expanded matrix and the sample matrix for later processing.

## Best Estimate Car Trip matrices for Base Assignment and Model Application

Non-residents' home-based car trips from the external cordon survey are not duplicated with any other data source and will be processed to give trip matrices by time period: these will be added to the residents' household survey trip matrices to give fully observed base matrices of all car travel.

Residents' home based trips from the external cordon survey duplicate the household survey, but it is likely that their sampling rate is much higher than the household survey. Therefore, they will replace the household survey data in the trip matrices.

[2] We can only combine the school and household survey bus matrices if we can isolate school trips from tertiary education trips (on which we have no supplementary data) in the household survey.

Non home based trips from the external cordon survey include both residents' and non-residents' trips and thus partially duplicate the household survey (which covers residents). The external cordon survey data will replace the household data.