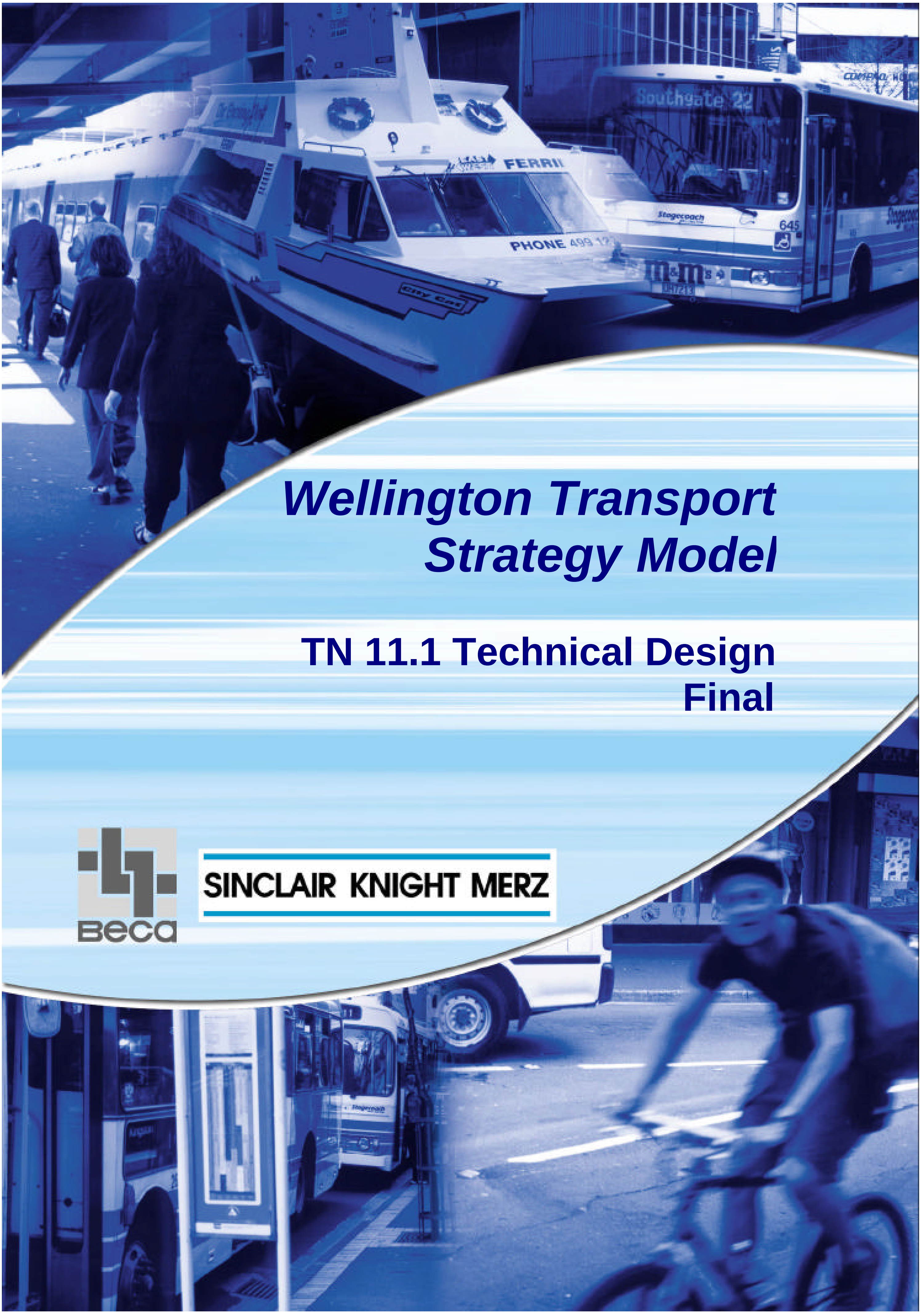


Wellington Transport Strategy Model

TN 11.1 Technical Design
Final

Beca

SINCLAIR KNIGHT MERZ

Wellington Transport Strategy Model

TN 11.1 Technical Design

Final

March 2002

prepared for

Wellington Regional Council

by

Beca Carter Hollings & Ferner Ltd

and

Sinclair Knight Merz

Contents

Document History and Status

Issue	Rev.	Issued To	Qty	Date	Reviewed	Approved

Printed:	1 August, 2003
Last Saved:	1 August, 2003
File Name:	P:\81\8110573\Task 11 - Technical Design\Technical Design (V6).Doc
Project Manager:	David Ashley
Name of Organisation:	Wellington Transport Strategy Model - Technical Design
Name of Project:	
Name of Document:	
Document Version:	
Project Number:	ME02030

1. Introduction

An effective means of controlling model developments is to pre-define as much of the work as possible, thus reducing later risks and reducing the likelihood of time being wasted on blind alleys etc. The purposes of the Technical Design are therefore to:

1) identify the full range of tasks needed to develop the model,
2) describe each task in as much detail as possible, in this process seeking to constrain the analysis to preferred/expected outcomes, and
3) anticipate risks.

This Technical Model Specification is written in note/abbreviated form in order that the key requirements are transparent. It is intended as a reference document for the model development activities rather than a model text and will be progressively extended and increased in detail as the project progresses.

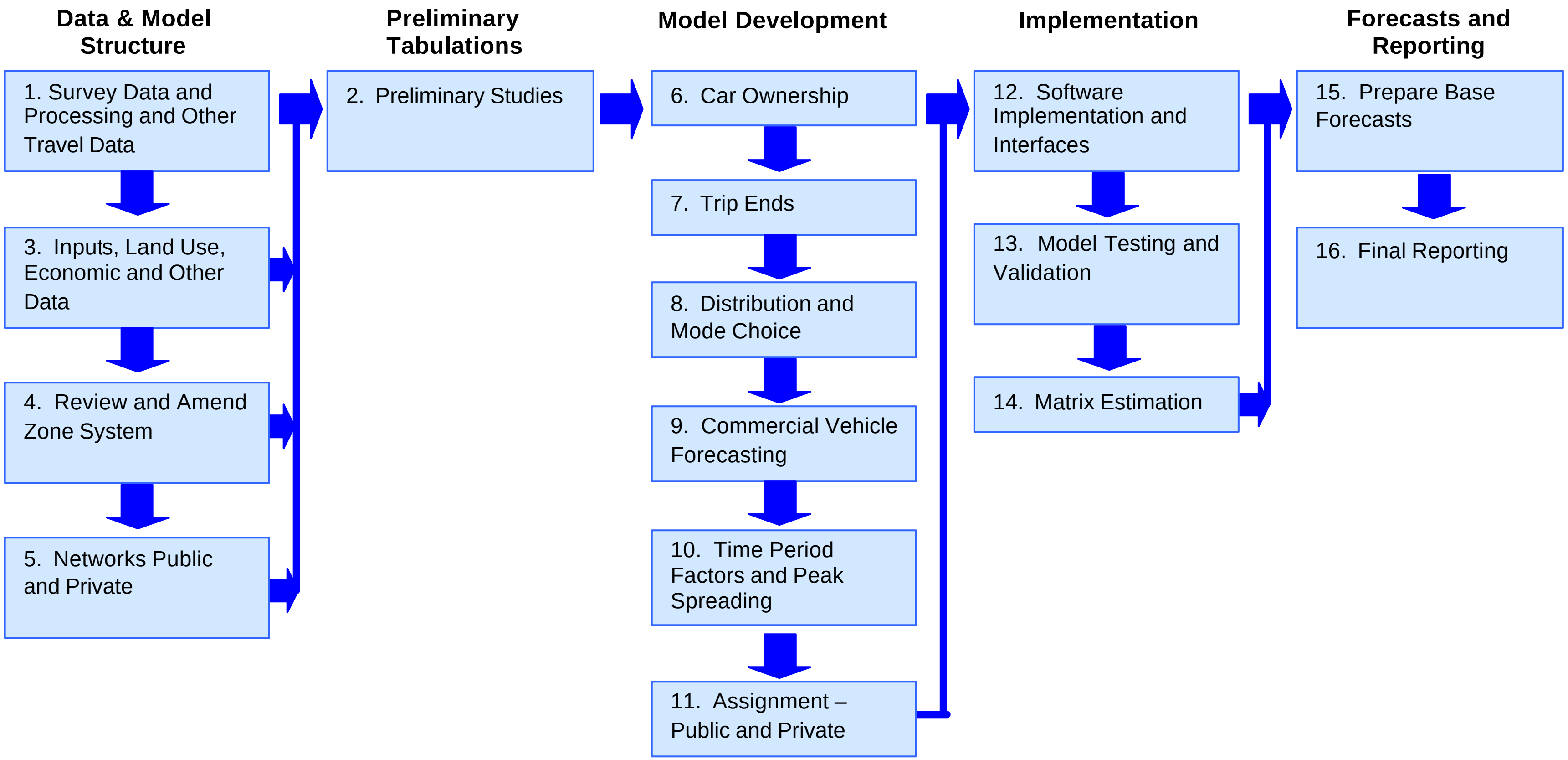
The main text is a series of tasks, structured with one task per page, with the format *Inputs - Processing - Outputs - Resources*[1]. In this version, the resource plan for each task budget is not completed, but should be reviewed in detail at the outset of each task. Additionally, where it is seen to be useful to detail sub-tasks, these are included as appendices.

The structure of the work programme is illustrated in Figure 1.1 and the individual tasks are presented in detail in Chapters 3 onwards, following some initial notes on model scope in Chapter 2.

For a project of this magnitude it is obviously not possible to anticipate every contingency at the outset. Therefore at the commencement of every task, the specification in this Technical Design should be comprehensively reviewed and, in some cases, further detailed.

[1] This convention is not followed in the later tasks whose dependencies on earlier work are such that a fully detailed specification has not been attempted.

■ **Figure 1-1 Task Sequence**



2. Model Scope

This chapter records the decisions so far taken on model dimensions as these effect all modules.

2.1 Constraints

Key constraints on the model development agreed at meeting on 8th May 2001 are

- a single-mode (road traffic) weekend model;
- a simplified approach to commercial vehicles (CVs) is to be used, further improvements being based on the Opus research for Transfund;
- specific developments in connection with road pricing and toll modelling are not required, although the model is intended to be capable of taking some account of these influences on travel behaviour[2];
- parking is only of interest for Wellington CBD, where the model should be able to forecast the impacts of parking supply and pricing strategies.

2.2 Model Segmentation

With linked models there is a need for consistency between segmentations adopted in each sub-model, and it must be feasible to estimate the population distributions for the chosen segmentations. The Functional Design and Appendix A of the Survey Strategy consider some of the segmentation options which we bring together here. In all models the key segmentation dimensions relate to: time of travel, trip purpose, person/household type and car availability.

The following attempts to pre-define a feasible segmentation for the whole model consistent with best practice but also recognising the limitations of the household survey sample. While it is not inflexible, and can be developed on the data, it should be recognised that extra segments may impose a disproportionate cost (to their incremental value) on model development.

2.2.1 Time periods

The model time periods will be as follows. It needs to be decided whether they are based on time of departure, arrival or midpoint of the trip, and whether this can vary by time period (eg am peak based on arrival time and pm peak on departure time).

[2] The implication of this is as follows. The model will be sensitive to prices and thus tolls and will include peak-spreading effects for differential peak and off-peak pricing. It also separates business travel from other purposes, which would have a different reaction to tolls. The model will thus permit the testing of toll strategies. Specific model developments for the specific purpose of improving the accuracy of toll modelling will not however be made: such as stated preference work on toll reactions, segmentation of travel by person/household income or specific attention to the particular problems of the very sensitive re-assignment effects for Transmission Gully.

Weekday: am peak: 7.00-8.59am interpeak: 9.00am-3.59pm pm peak: 4pm-5.59pm offpeak: 6pm-6.59am Weekend: an average of Saturday and Sunday daily traffic Saturday morning peak period: to be defined (eg 12.00-14.00) Sunday evening peak period: to be defined (eg 12.00-14.00)

2.2.2 Trip Purpose

As discussed in Appendix A of the Functional Design, the following trip purposes will be used.

Personal travel: ❑ HB work – HBW ❑ HB education – HBEd ❑ HB shop – HBSh ❑ HB social – HBSo ❑ NHB other – NHBO	Applies to sub-models: ❑ trip ends ❑ mode choice ❑ distribution ❑ peak-spreading ❑ time period factors
HB/NHB Business travel – BU	
Travel by commercial vehicle (sub-segmentation to be determined) – CV	

To be defined:

- ❑ the segment in which to include HB Other (shop, social or NHB Other),
- ❑ the treatment of escort/serve passenger trips
- ❑ the vehicle types which are CVs

2.2.3 Car Availability: Choice/Captive

The general principle of splitting the market into choice and captive segments for mode choice modelling used in the present model will be retained, with a possible refinement as follows.

Captive	trips by residents of non car owning households	Applies to sub-models: ❑ trip productions ❑ mode choice ❑ peak-spreading time period factors
Competition	trips by residents of households where no. of cars < no. of adults	
Choice	trips by residents of households where no. of cars in household ≥ no. of adults	

Notes.

(i) For a pre-distribution mode split structure, the captive choice segmentation would not be used in the distribution model.

(ii) A simplification has been made by using a somewhat weaker definition of competition for the car, whereby the comparison is made between the number of cars and the number of **adults** rather than licensed drivers. This avoids the additional burden of having to model licensed drivers explicitly.

(iii) Definition of adult is to be specified, given driving age is 15.

Preliminary checks suggest that the split of households between the captive, competition and choice segments may be broadly 10%, 20% and 70% respectively, providing some justification for the additional competition segment (these statistics need to be verified).

2.2.4 Person/Family Structure

These requirements primarily affect the car ownership and trip end sub-models. The present family structure model is unsatisfactory in that the 8-way household category is generally ambiguous regarding the number of persons in the household, a major determinant of the number of trips.

The MERA demographic model provides number of persons by age and sex categories by zone and the number of households according to the 18-way categorisation.

Allowing for what has been thought important in Wellington, in other NZ models and in London, the following dimensions to the categorisation are proposed for the updated WTSM (and this will need to be confirmed in preliminary analyses).

Trip Production Model Segmentation	Source
Person types (age classes to be confirmed): ❑ infant (<5) ❑ child (5-16?) ❑ young adult (17-25) unemployed ❑ young adult (17-25) employed by: – full time – part time/casual – other (self-employed etc) ❑ adult (26-60) unemployed ❑ adult (26-60) employed by: – full time – part time/casual – other (self-employed etc) ❑ "retired" (>60)	Classification by age: as MERA. Classification by employment status: as MERA or age-specific activity rates. Classification by employment type: fixed proportions from census, amended for forecasts as scenario parameter.
Household characteristics: ❑ car ownership: 0, 1, 2, 3+ ❑ possible other effects	Car ownership from relevant sub-model.

Notes:
(i) Expectation is that the 3+ car segment may be about 9% of households in 2001.
(ii) While income is an interesting explanatory variable for discretionary travel, it is likely to be too highly correlated with car ownership to be included in the models.
(iii) Household size may be an influence on person trip rates for some purposes, though it is probably an unnecessarily demanding refinement.
(iv) It is not assumed that the "retired" category does not commute.

The trip production model will be based on person trip rates according to the following person types. The trip production models may also include the 'main effects' of some household characteristics. One reason for the detail of the car ownership model is that non-work person trip rates are expected to be higher for the more mobile car owning households.

As a separate but related issue, the estimated trip productions must be classified into captive, competition and choice trips for the purposes of predicting mode shares, and the main information for this is derived from the car ownership model.

The car ownership model will also be segmented by household characteristics.

Car Ownership Model Segmentation	Source
No. of adults: 1, 2, 3+ Employment status: retired, non-retired	Classification by adults & employment status: as MERA.

Notes:
(i) Other parameters such as presence of employed persons and children in the household have also been included in other models, though again this seems a demanding refinement.
(ii) Makes 5 household categories: retired: 1, 2+ adults; non-retired: 1, 2, 3+ adults.
(iii) It is to be decided whether this is simply an age classification (adults > 60) or an employment status classification, with the former being preferred.

2.2.5 Implementing the Proposed Segmentation

In order to implement the above segments in the model changes must be made to the MERA land use model or supplementary methods introduced as part of the transport model.

The MERA Model

The MERA model projects zonal population by age and sex. The allocation to the present household types is done mechanistically from these basic projections. Although the process is scarcely described, it appears to be the following:

- current census data is used to establish for each person type the probabilities of being in each household type;
- given then the populations in each household type, an average household size determined from the census was used to determine the number of households of each type; the extent of disaggregation of the average household size estimates in not clear;
- the model specifically does not allow for any future trend in household size;
- the employed/unemployed distinctions are introduced apparently using the above procedure rather than any explicit activity rate (by age and sex) implementation.

One part of the review of the process will be concerned with the value of this second stage of the MERA model.

Review of The New Requirements

The position is as follows:

- the person type segmentation for the trip production model should be available through MERA; if the employment status is not, then it could be readily addressed using age and sex activity rates;
- the household type segmentation for the car ownership model: the household categories used in the present model are distinguished (among other things) by

number of adults and employment status; the precise definitions are probably not the same, but it is clearly no great task to amend them.

The key issue is how to link the car ownership and trip production models to produce the forecasts of trip-making by car availability and person type.

Car Ownership Model Segmentation			Trip end Model Person Types P1	P2	P3	P4	P5	P6	P7
1	retired	1 adult							
2	non-retired	1 adult							
3	retired	2 adults							
4	non-retired	2 adults							
5	"	3+adults							

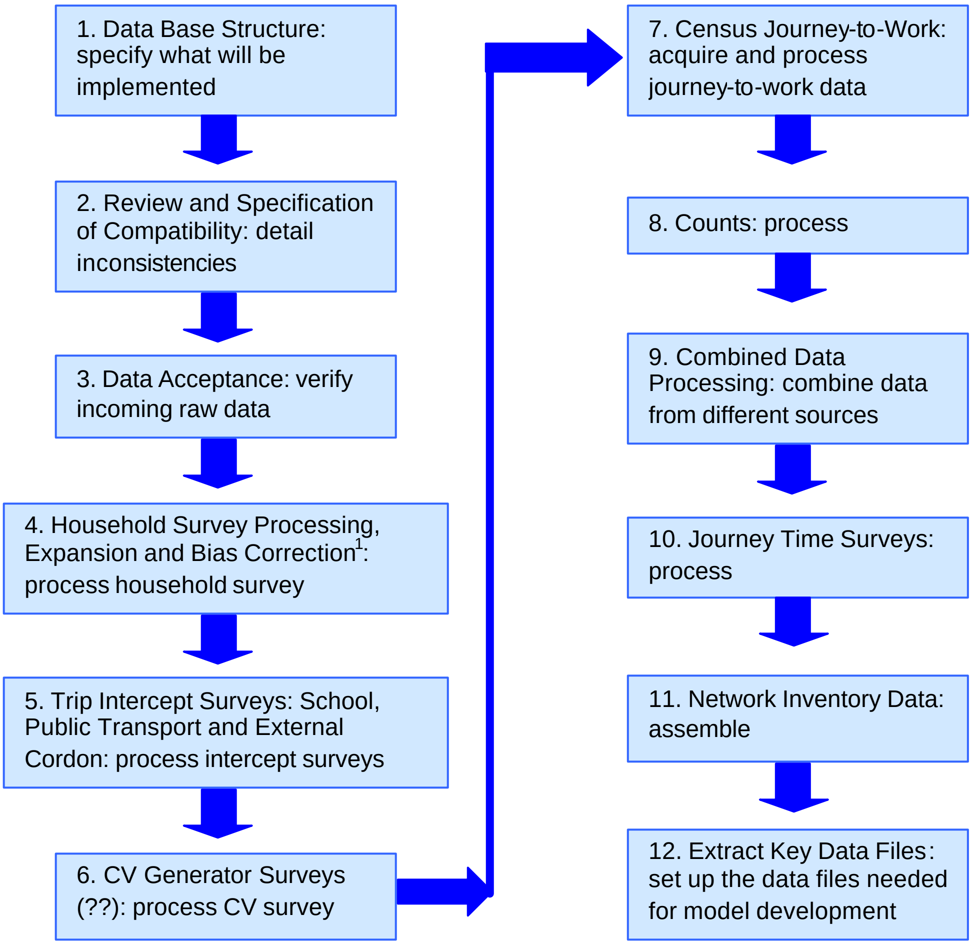
This must be further refined for the car ownership segmentation into the categories illustrated below. The procedures involve, for each household category, allocating persons of each type between the household sub-categories of car ownership. A procedure for this is discussed in the section on family structure.

Car Ownership Model Segmentation				Trip End Model Person Types P1	P2	P3	P4	P5	P6	P7	Car Availability
1.1	retired	1 adult	0 cars								Captive
1.2	"	"	1 car								Choice
2.1	non-retired	1 adult	0 cars								Captive
2.2	"	"	1 car								Choice
3.1	retired	2 adults	0 cars								Captive
3.2	"	"	1 car								Competition
3.3	"	"	2+ cars								Choice
4.1	non-retired	2 adults	0 cars								Captive
4.2	"	"	1 car								Competition
4.3	"	"	2+ cars								Choice
5.1		3+ adults	0 cars								Captive
5.2		"	1 car								Competition
5.3		"	2 cars								Competition
5.4		"	3+ cars								Choice

3. Task 1 Survey Data and Processing and Other Travel Data

3.1 Scope

This task has the purpose of creating the data files needed for model calibration and implementation from the raw survey information. The sub-tasks are as follows.



3.2 Task 1.1 Data Base Structure

Inputs
Survey Specifications and coding frames.

Processing
Design data base structure providing for convenient data storage and processing – mainly concerned with allowing for hierarchical household survey structure and facilitating its conversion into flat files as needed.

Outputs
Specification of data base.
Report.

Resources

3.3 Task 1.2 Review and Specification of Compatibility

Inputs

Questionnaires, and coding frames/instructions for all travel data:

- household survey
- public transport survey
- external survey
- school survey
- census journey to work

Processing

Particularly affecting data not specifically assembled for this project (eg the census), a specification of the consistency of definitions for key variables is required plus an equivalence table, covering at least:

- zoning,
- main mode of transport for a trip,
- cars
- household size/present/absent on travel night,
- employment status
- car ownership/availability,
- purpose,
- definition of surveyed trip (eg travel day, typical day ...)

etc

Outputs

Equivalence tables.

Report documenting the consistency of definitions for key variables.

Resources

3.4 Task 1.3 Data Acceptance

Inputs
Survey data files provided by subcontractor(s) for household, public transport and school surveys.

Processing
Pre-specify acceptance tests: specifically these must identify all MIC issues.
Install data in data base.
Undertake acceptance tests.

Outputs
Report documenting acceptance test results and overall conclusions.

Resources

3.5 Task 1.4 Household Survey Processing, Expansion and Bias Correction

Inputs
Household survey data file and coding specification.
Base year land use data (census households by type by zone).
MIC Specification.

Processing
1. Specify census data required for expansion purposes and order
2. Implement common purpose codes (Appendix 1.1).
3. Deal with MIC issues (Appendix 1.2).
4. Trip linking (Appendix 1.3).
5. Expansion and bias correction (Appendix 1.4).
6. Annualisation (Appendix 1.5).
7. Use of extra PT trip data[3].

Outputs
Trip-linked, expanded household survey data file.
Technical Note.

Resources

[3] To be specified.

3.6 Task 1.5 Trip Intercept Surveys: School, Public Transport and External Cordon

Inputs
Interview data bases.
Count data bases – *to be specified*.

Processing
Main stages/tasks, for which specifications are required:
- implement common purpose codes (Addendum 1.1)
- expansion and annualisation:
  - expansion of each base market survey to travel in the interview period (eg Monday, 6am-8pm);
  - expansion from interview to survey period (eg 1 or 2 full weeks);
  - expansion to annual total (Appendix 1.5);
  - elimination of out-of-scope data and double-counting;
  - elimination of known biases;
  - verification of expansion.
- trip reversal.

External Roadside Survey – see Appendix 1.5
Rail and bus survey – see Appendix 1.6
Schools survey – see Appendix 1.7

Outputs
Expanded data base of intercept surveys.
Report.

Resources

3.7 Task 1.6 CV Generator Surveys

It is suggested that the cost and benefits of selected cordon counts around specific generators such as the ports and airport should be investigated.

3.8 Task 1.7 Census Journey-to-Work

Inputs
2001 census journey to work data.

Processing
Prepare a specification of requirements:

- precise data requirement,
- receipt and checking of data,
- processing into suitable form, including whatever scaling may prove necessary to represent an average day: there are significant issues of definitional inconsistency with what is required for WTSM; these must be identified and so far as possible eliminated:
  - compare questionnaire definitions in census and household survey; note differences;
  - compare survey context and note differences (issues include: census ignores details of journey-to-work, for example 'tours' involving dropping off children at school, shopping, going out after work ....;
  - specify how to reconcile household and census data

Order data.
Implement processing.

Outputs
Census journey to work matrices by mode and sub-segment.

Resources

3.9 Task 1.8 Counts

While traffic counts would be available through WRC, it is expected that these will need to be supplemented to ensure adequate coverage and to enable accurate vehicle classification. Equivalent data for public transport passengers, possibly taken from ETM data, will also be required.

Inputs
Road (classified)
Manual classification counts, including an attempt to identify vehicle type for commercial-purpose vehicles?
ETM data?
Other?

Processing
<u>Road traffic counts</u>
The existing WRC traffic count data will be reviewed for completeness for our purposes (screenline coverage, ability to distinguish vehicle types). A programme to rectify omissions will then be designed and executed.

If a commercial vehicle model is to be developed it will be necessary that these counts are classifiable by vehicle type.

The complete set of data must be checked for completeness, reliability etc, processed into a convenient form and verified

Key outputs:
- Database of links flows by modelled period
- Typical daily/weekly flow profile (to verify time period selection, develop assignment peaking factors etc
- Daily flow factors (factors to estimate daily link flows from individual time period flows)

Investigate flow profiles from a representative selection of traffic counts. Investigate temporal variations by area. Verify suggested model time periods against link flows.

Develop ‘peaking factors’ for am and pm peaks. Requires count data at 15-minute intervals. Calculate ratio of average flow for all 15-minute periods greater than the average flow, to the average flow. Peaking profiles should be developed from single-day counts rather than from a profile averaged across a number of days. This is to avoid the daily variations smoothing the profile.

Create database of link flows. Each site referenced by model IJ nodes, count date (year and week) and including period totals and 5-day and 7-day totals.

Investigate historical trends where historic data available. Generate global and/or area-specific growth rates. ‘Normalise’ non-2001 data to equivalent 2001 flows using appropriate growth rates. Investigate seasonal variation, and if possible create seasonal factor table by count week.

Develop ADT factors for a range of peak models, e.g. factors to estimate daily flows from following models: am and interpeak only; am, interpeak and pm

Public Transport Passenger Data

The feasible direction for this is to be established (ETM, special counts, screenline observations).

Outputs

Count data base
Typical and area flow profiles
Period peaking factors
ADT factors

Resources

3.10 Task 1.9 Combined Data Processing

Inputs
Trip matrices from household, school and public transport surveys and from the journey to work census.

Processing
The general objective is to make the best use of these data in model development and application. Until the surveys are well-defined, we cannot make a great deal of progress, other than enunciate the principles.

In model calibration, the data generally need to be taken from one consistent source and be available to the full level of model segmentation – a requirement which generally favours basing model estimation on the household survey (*this needs reviewing in the context of each model calibration exercise).*

However for model application, the model trip matrices may be substantially improved by drawing on trip patterns derived from surveys with higher sampling rates, and in this case it may not be necessary to have available to the full model segmentation, some degree of aggregation being quite acceptable.

Whatever the application, there are two stages to combining data into trip matrices:
- the first is to verify some general data compatibility, that is a consistency in data aggregates like total trips, average trip length etc etc;
- given consistency, then the second is to combine the data using an inverse variance approach (and we probably know enough about the data variance to address this)[4].

This approach could be used for HBW (which we may simplify to basing the best estimate on the census, given the disparity of sampling rates), HBEd, for school-level travel only and public transport trips in the northern corridors for all purposes. The matrix below summarises the sources for various travel segments.

Purpose/Mode	**Car**	**Public Transport**	**Other Modes**
HBW	Household Census	Household Census Public Transport	Household Census
HBEd	Household School	Household School Public Transport	Household School
Other purposes	Household	Household Public Transport	Household
Resident external	Household External roadside	n/a	na/

Tasks involved would be:
- specify basis of each data source and their mutual compatibility (concerns: method and coverage of sampling and consistency of definitions – purpose, mode etc)[5];

[4] The general formula for comb ining trips (T) for segment s(and zones i to j from two data sets with variance (V) is:
$T = (V_1 * T_2 + V_2 * T_1)/(V_1 + V_2)$ (*check*)

- produce tabulations comparing estimates from each source
- specify which sources should be combined and the statistical process
- implement.

A final requirement is to generate suitable vehicle matrices for a 2001 assignment, from which journey times and costs would be extracted for model calibration. Observed matrices derived from the household survey are likely to be very sparse, so there is some value in considering a combined observed and synthesised (from the old model) matrix, scaled using matrix estimation techniques to counts.

Outputs
Best estimate base trip matrices.
Note.

Resources

[5] Concern over the census data that it does not allow for linked trips/tours, eg dropping children off at school en route to work and this inconsistency (and others) will need to be adjusted for.

3.11 Task 1.10 Journey Time Surveys

3.11.1 Survey Specification

Inputs
Selected routes to be determine as part of validation criteria.

Processing
Journey time surveys to be undertaken on between Tuesday and Thursday.

Each route is to be surveyed every 15 minutes through the peak AM and PM two-hour periods. Eight survey runs are to be recorded for the AM and PM peak periods.

For the interpeak every hour is to be surveys between 9:00 – 16:00.

The Saturday and Sunday peak two hour model is to be surveyed similar to the AM and PM peak periods.

Derive the mean journey time and journey times to 2 standard deviations for each route and peak period based on survey data recorded.

Outputs
Journey time statistic required for validation

Resources
Resources and budget have been allocated as part of other surveys.

3.11.2 Processing

Inputs
Journey time surveys (road)
ETM data for PT times?

Processing
Objective is to compare cumulative time graphs ('worm diagrams') on complete survey routes, as well as average times between landmarks (sub-sections of route).

Process road survey data to extract cumulative travel time along length of route, obtaining average, minimum and maximum survey runs.

Sub-divide survey routes into feasible sub-sections and obtain min, max, average and standard deviation times for sub-sections. Sub-sections should not be too short or too long – preferable between 4 and 10 minutes.

Develop interface with EMME/2 network coding: Suggest use of a link tagging system which can identify both the overall survey route and the sequence of links within the route (Link data must be tagged so that it can be sorted by chainnage along route)
Create macros for extracting relevant data from EMME/2
Develop spreadsheet to compare model and survey data, both cumulatively ('worm diagrams'), and for point-point totals on sub-routes.

(Note: these procedures, macros and spreadsheets have been developed on other projects and hence only need to be adapted slightly to work here)

Process a sample of ETM data and compare with timetables where possible (eg in Newlands/Mana) or otherwise with direct observation.

Outputs

- Min, max and average cumulative travel times on each surveyed route
- Interface between model and data (emme/2 macro using
- Validation spreadsheet comparing model and observed time data – both cumulative and point-to-point.

Resources

3.12 Task 1.11 Network Inventory Data

Note that the proposals assume that no specific inventory data collection is required other than speed surveys.

Inputs
GIS database – census meshblocks, road centrelines, TLA boundaries, traffic zone boundaries
District Plan hierarchy maps
Transit NZ route data sheets (State Highway only)
Street maps
Sample free-speed surveys (with sample from each link-type classification)
Inventory surveys at key junctions
TLA data on clearways

Processing
Basic task is to develop a robust link-type classification, identifying:
TLA
RCA
Road environment (service road, local road, arterial, rural, expressway etc etc)
Posted speed-limit

This classification will form the basis for the initial allocation of link-free-speeds, capacities and vdf, and the generic relationship between link type/speed limit and the free-speed. Free-speeds by link-type should be based on sample free-speed surveys.

Network data that varies by time period should be collated (clearways, bus lanes, tidal-flow facilities), and identified by link i-j reference. This data is held by TLAs/RCAs.

A library of the layout details for all key junctions should be developed. This should include a sketch of the layout, including lanes, lane markings, basic lane dimensions (widths, short lanes etc), phase diagrams etc. These should be sketched on pro-forma forms (then scanned into electronic format along with digital photos??). Develop an electronic storage system for junction library.

Outputs
Link-type classification of all links
RCA/TLA/DP Hierarchy classification of all links
Generic allocations of free-speeds and capacities by link type
Junction layout library

Budget

3.13 Task 1.12 Extract Key Data Files

Inputs
Data Base.

Processing
A very large number of data files (eg car ownership, trip ends, trip matrices, generalised costs) needs to be extracted on a fully consistent basis. The required files must be specified and fail-safe procedures need to be set up and the files extracted.

Outputs
Analysis files.

Resources

4. Task 2 Preliminary Studies

4.1 Scope

These are kick-off tasks designed to confirm the early decisions in Chapter 2 and to clear uncertainties concerning other aspects of the overall model or its structure.

1. Review of performance of present model: to ensure that we understand its weaknesses and, so far as practicable, have plans for overcoming them
2. Initial Tabulations: designed to confirm the segmentation decisions
3. Analysis of parking data: to firm up on the treatment of parking in the model
4. Specification of Generalised Cost: on which the model calibration will be based
5. Retail Destination Analysis: to develop the treatment of retail destinations in the model
6. Commercial Travel: to confirm the approach to commercial travel
7. Education Modelling & School Buses: to review the modelling of education trips and the representation of school buses
8. Car Passenger Modelling and Escorts: to determine the treatment of escorts in the model
9. Weekend travel: to design the weekend travel methodology
10. Road Pricing/Tolling
11. Commercial Vehicle Travel
12. Use of Intercept Matrix Data
13. Vehicle Types
14. Model Structure Simplifications
15. Park-&-Ride
16. Ports and Airports

4.2 Task 2.1 Review of Performance of Present Model

Inputs
Previous WRC study of this issue.
Additional historic volume data split into peak and interpeak periods.
Historic demographic and economic information.
Current WTSM forecasts from 1991-2001.

Processing
Objective is to complete the WRC analysis of model forecasting performance, which compares historic traffic forecast between 1991 and 2001 with model forecasts. At present this task is probably mainly to extend the current daily analysis to the peaks and interpeak periods where trends might be expected to differ.

Extent of the work and its feasibility needs to be reviewed and a more detailed specification prepared.

Outputs
An assessment of actual travel trends in key corridors in different time periods against the main explanatory factors (demographic and economic).
An assessment of the model's performance in reproducing these changes.
Conclusions as to model specification developments which might improve the representation of these trends.
Report.

4.3 Task 2.2 Initial Tabulations, Model Structure and Segmentation Decisions

Inputs
Survey data base.
Other data sources as required.

Processing
Objective is to confirm model structure decisions and precise definitions of model segments (including CV vehicles, age classifications, definition of adult etc).

Specify required tabulations.

Implement tabulations:

- tabulate household survey to verify purpose and captive/choice[6] segmentations[7] (diagnostics: number of trips, mode shares, trip lengths and % to CBD);
- tabulate household survey and volume counts (PT counts/ETM data) to verify choice of time periods, including at the weekend;
- other

Outputs
Conclusions on model structure and segments.
Note.

Resources

[6] The significant issue is that captive was typically only 10% of the journeys for each purpose in 1991, which is very little data on which to base a model and accounts for quite a small proportion of total travel. However, this segment may account for a high proportion of public transport users.
[7] Including the 3+ car segmentation.

4.4 Task 2.3 Analysis of Parking Data

Inputs
Household survey.
Parking supply and pricing data (WRC).

Processing
Objective is to firm up on model treatment of parking and establish generalised cost parameters for parking.

Establish:
- for CBD destinations, tabulate by trip purpose the parking demands by type (long/short), location, type of car park and payment from household survey;
- obtain CBD parking supply data by type (long/short, park type); location?
- obtain the price of parking by parking type; location?

Parking supply and pricing data should be available from WRC.

Construct a model of average parking price for each trip purpose and of the total parking supply/capacity by long/short. Consider whether this has model application.

Outputs
Park price to be used in generalised costs.
Parking supply data, with a view as to its application in the model.
Note.

Resources

4.5 Task 2.4 Specification of Generalised Cost

Inputs
Present model.
Transfund present values and future changes.

Processing
Specify generalised costs by mode to be used in the model including coefficients. Eg:
- Car driver/passenger: time, direct operating cost, parking and tolls
- Bus/train: in vehicle time, access/egress time, waiting time, fare, interchange
- Walk/cycle: time

Outputs
Generalised cost specification.
Note.

Resources

4.6 Task 2.5 Retail Destination Analysis

Inputs
Household survey.
Planning data, including shopping centre classifications.

Processing
We currently anticipate that the most we might do is to include a dummy variable in the trip attraction model for HB Shopping identifying whether the zone contains various scales of shopping centre (strip shops, malls, bulk retail, CBD shopping areas) – in case the major centres attract higher trip rates. This analysis would be part of the trip attraction model calibration.

We might however consider some limited pre-analysis of the data to gain a better understanding of the characteristics of trips to retail destinations ("shops/mall/retail" places in the household survey. We might look at the geographic distribution of shopping trips, at the recreational travel to shopping places, at mode usage to/from and between shops, and we might try and relate this to a classification of zones according to the nature of the shopping available. The value of this needs to be reviewed.

Consult Michael Douglas.

Outputs
Note.

Resources

4.7 Task 2.6 Commercial Travel

Inputs
Household survey.
Planning data.

Processing
We have identified all business travel as a separate model purpose. Otherwise, there may be no specific model features.

Like retail travel, we may choose to do some pre-analysis of business travel to establish its major characteristics.

We shall need to determine overlap with CV model (relating to the vehicle types included in this model).

Outputs
Note.

Resources

4.8 Task 2.7 Education Modelling & School Buses

Important Note
WRC do not consider this task to be a priority. It will therefore only be implemented if WRC gives formal approval on the basis of further discussion of its value and scope.

WRC's view is that "little is achieved in this area in other models". We concur with this. But our concern is that most present models are incorrectly specified and likely to give wrong forecasts in a context of increasing importance – the increasing use of cars to take children to schools and the impacts this has on peak congestion.

It is not however clear whether there are practical ways of improving on conventional modelling approaches and this task is designed to review this before the modelling commences.

Inputs
Household survey.
School survey initial analysis.

Processing
The conventional model structure appears to offer little for education trip modelling: trip distribution is very constrained by discrete school locations and involves comparatively short trips while children by definition cannot drive a car but may have wider car availability through parents and lifts. The treatment of school bus travel in the model needs also to be considered given that such services are not included in the networks.

The purpose of this task is to carry out some analyses of the household survey as an aid to reviewing the model structure. It will be important to consider the expected trends and the policy issues.

Recognising that little is achieved in this area in other models, this task may well not lead to any change in model specification.

Some preliminary thoughts which could be followed up:
- key trip characteristics (eg mode, trip length) may be a function of education level (primary, secondary, tertiary), as children get more independently mobile with age (although it is not likely that we would contemplate an additional segmentation in the model);
- we are particularly interested in the choice of mode between walk, public transport, school bus, family car passenger and other car passenger and how this varies by education level, distance and, perhaps, family characteristics.

Outputs
Note.

Resources

4.9 Task 2.8 Car Passenger Modelling and Escorts

Important Note
WRC do not consider this task to be a priority. It will therefore only be implemented if WRC gives formal approval on the basis of further discussion of its value and scope.

WRC's view is that "little is achieved in this area in other models". Our concern is that practice varies around the world and that this particular segment appears in some data sets to account for a high proportion of travel. If this turns out to be the true of the 2001 Wellington household surveys, then it would seem that some attention should be given to this segment. However, if escort trips account for a small proportion of travel, then we would not consider it to be an issue.

Inputs
Household survey.

Processing
Car pooling issues seem to be of most interest for work and education trips[8]. Mode choice models either treat car passenger as a separate mode or factor out such trips on assumed car occupancies. Neither approach appears to replicate the likely choice behaviour, but research into this area suggests very substantial model complications to represent the real behaviour.

The purpose of this task is to carry out some analyses of the household survey as an aid to considering the most appropriate model structures. It will be important to consider the trends which might be expected (in relation to historic trends) and the policy issues. Perhaps the major decision to be made is whether escort trips are to be combined in some way with other trip purposes or separately modelled.

There is some overlap with Task 2.7.

Recognising that little is achieved in this area in other models, this task may well not lead to any change.

The following are some ideas:

- the principal analysis is to understand the varying nature of the escort trips: we might for example expect most to be as a car driver for the purpose of carrying children of the family on school, social or recreational activities; this will involve analyses which relate the car escort journey to journeys by other family members;
- we certainly want to differentiate escort trips which are part of a 'tour' (eg father en route to work dropping off child) from those which are solely escort generated;
- if we can understand and classify such journeys we may be in a better position to consider how they should be forecast and in relation to what explanatory variables: at one extreme, we might forecast the car passenger trips of children in the mode choice model and then automatically generate the associated car driver trip, and this would be the education escort model;

[8] For other personal and business trips it would be our view that car occupancy factors would be sufficient to represent household car sharing and car sharing on business.

- a key issue is to reflect in the forecasts what we expect to happen if the passenger (the cause of the trip) is forecast to change mode (expecting that the driver's trip will disappear).

Outputs
Note.

Resources

4.10 Task 2.9 Weekend Travel

Concepts
The objective is for a model of weekend travel whose main output is a road traffic matrix and assignment. We expect to gain about 1,000 weekend household interviews spread between Saturday and Sunday on which to base the model (about 40% of the weekday sample). We are not aware of any city models with a weekend implementation (which does not necessarily mean that there are none)[9]. We thus have little prior information/precedents on which to base the model specification, so this task is concerned with confirming the specification proposed below.

The demand models will be based on the travel for the whole weekend (ie Saturday and Sunday), with subsequent time period factors generating any sub-period. This is a strong hypothesis which presumes that, while the purpose mix can be different between the 2 weekend days, the patterns of travel for any given purpose are comparable. This will need to be checked in initial tabulations. But it is probably an unavoidable assumption, the interviews for a single day being insufficient for model development.

A major issue is what if any consideration is to be given to public transport trips. With such a small household sample and the (presumed) relatively light use of weekend public transport, there will be few such trips in the data and the public transport share will be low. For these reasons, the proposed model includes public transport trips at the trip end stage then uses fixed mode shares to identify car-based trips, which are then distributed and assigned. Such a structure could later be extended to include public transport if worthwhile. A great advantage of this approach is that it is unnecessary to set up a weekend public transport network, a potentially substantial task.

<u>Road Network.</u>
A further important issue concerns the road network representation and the stability of travel profiles through the day. If an <u>average</u> hourly trip matrix (traffic level) for the weekend days is used then we would expect most of the network to be uncongested. However, in the weekend peak periods, we expect significant congestion in some parts of the network, although these are not necessarily the same on Saturday and Sunday. In order to simplify the modelling, we propose to generate the weekend trip matrices using this concept of an average hourly matrix, but to create assignments for the peak periods by factoring this matrix appropriately.

<u>Car ownership and family structure models.</u>
Obviously, these apply equally to the weekday and weekend.

<u>All Mode Trip Ends.</u>
We suppose that the dominant purposes will be social and recreational, implying the main trip purposes would be HBSh, HBSo & NHBO. There will be some work, education and business trips, but we expect them to be of much less significance. Before we can determine whether these should be individually segmented or

[9] David Ashley, while with MVA, developed the Long Distance Travel Model for the Dutch Government which included representation of weekend interurban travel.

amalgamated in some way we shall need tables to establish their significance through survey tabulation. Despite their potentially minor significance, consistency with the weekday models may mean it is convenient to retain an individual segmentation.

The trip production models will be of the same structure as those for the weekday. But the coefficients (trip rates) will be different. It will probably be worth looking at joint estimation of the weekday/weekend combined data set – if the effects of person and household type on trip rates are similar between the weekday and weekend, this will enable the weekend model to profit from the larger weekday data sample.

A similar approach will be taken with trip attractions. But it is conceivable that the model parameters could be different (the clearest example is work trips, where at the weekend these seem likely to be predominantly by retail employees, implying a different model structure to the weekday).

<u>Mode Choice/Distribution.</u>
The structure proposed for the weekday will be used (Figure 10.1), except that the right hand side of this figure concerned with modelling public transport and slow modes will not be implemented. Instead, the 2001 car mode shares for choice trips by zone and segment will be applied as fixed factors in the model in future forecasts. The distribution model will be for car trips alone. A joint estimation with weekday does not seem feasible, but the weekday parameters will provide a startpoint for the weekend model estimation.

<u>Commercial vehicle forecasting.</u>
To be discussed separately, but presumed to be less important at the weekend.

<u>Peak-Spreading</u>
Not applicable. We note that there are issues relating to the stability of time profiles across the weekend – how many trips are restricted in their choice of time of travel and how many have a free choice? These are not matters we intend to address in WTSM.

<u>Assignment.</u>
In principle, identical to weekday. Minor point of difference:
- network differences in one-way systems etc, although none are expected;
- different speeds.

<u>Time Period factors.</u>
Time period factors are required for two purposes:
- choice of 'average' hour to use for assignment;
- development of flow forecasts for specific weekend periods.

This is an area of some significant uncertainty. The following are some of the issues:
- there may be key time periods for policy reasons, for which we need client advice; two examples are suggested (times to be confirmed):
  - Saturday: 12.00-14.00 (a morning shopping/recreational peak)
  - Sunday: 12.00-14.00 (although there may be an evening return home peak[10]).

[10] In practice, if as seems likely the Sunday evening trips for each purpose are not typical of those at other times of the weekend, then

- the time periods need to be distinctive in terms of the mix of purposes and the direction of travel if the time period factors are to be effective;
- in principle, although we are not aware of this being done elsewhere, the time period factors could be disaggregated by parameters other than purpose and travel direction including, for example, trip length.

Matrix Estimation.
The final stage will be to fit the vehicle matrix to count data for the specified periods of interest.

Risk and Uncertainty
Unlike much of the weekday model, the development of the weekend model takes us into untrod territory, and therefore involves an element of research and some risk concerning the adequacy of the outcomes. Such technical risks add to costing and resource uncertainties.

Inputs
Household survey.
Counts.

Processing
The purpose is to confirm the weekend model specification through tabulation of the household survey and analysis of weekend traffic count data.

Outputs
Specification.

Resources

this method of estimating them is likely to be poor. This is another reason for the recommendation to adopt the 2 midday periods for the peaks.

4.11 Task 2.10 Road Pricing/Tolling

The current model structure is in principle capable of providing broad estimates of tolling and road user charging impacts. But it does not include refinements associated with tolling models (segmentation by the income of the traveller, sophisticated peak-spreading models, willingness-to-pay surveys etc). There may also be special issues for projects such as Transmission Gully (where the context is such that the sensitivity of the evaluations to assignment techniques and reaction to tolls may be unusually high).

Tolling applications may be at the strategic level involving the use of WTSM directly (for strategy appraisal) or at the project level (such as Transmission Gully) where a project model, probably developed from WTSM and using some WTSM inputs would be used. There are thus two ways in which WTSM might be used: directly or indirectly in relation to project models.

Because tolling issues are of increasing importance, there is a need to describe how the proposed, updated WTSM would be used in tolling studies, allowing also for the project model interface. This will both clarify how such studies will be undertaken and allow WRC to take a view on the acceptability and reliability of the procedures. If unacceptable, refinements to the WTSM specification will then be considered. This review will be entitled 'A review of the application of tolling and pricing in the Wellington Transport Strategy Model'.

In this review, it will helpful to focus on illustrative potential projects/issues, which WRC will supply.

The task will thus document:

- how the WTSM would be applied to strategic tolling studies, and its capabilities and deficiencies in this regard;
- how it would provide inputs to project studies and what additional features may be required in project models; again the capabilities and deficiencies would be addressed;
- how these representations might be improved based on international practice (a sort of shopping list).

4.12 Task 2.11 Discussion of Commercial Vehicle Modelling for the Updated WTSM

Introduction

Various ideas on how CV modelling might be approached were put forward in TN3.1 the WTSM Functional Design (FD). These were not developed in the Technical Design in the expectation that some of the issues could be resolved by the Transfund research project on Commercial Vehicle Usage and Forecasting. Unfortunately, this is not the case and this section therefore discusses the FD options. If the approach is agreed, the later related tasks will be detailed.

The note is written from two principle perspectives:
- the importance of heavy CVs as policy priorities (Table 2-4, FD);
- and WRC's view that heavy investment in CV modelling could be deferred to later model phases.

General Approach

As raised in the FD, a relatively inexpensive interim approach would comprise:
- development of rough 'prior' CV trip matrices;
- collection of CV screenline count data;
- relating the prior matrix to the current counts through matrix estimation;
- a growth factor procedure;
- an assignment procedure.

It is this approach which we discuss herein.

Vehicle Types

Difficulty of identification hampers most CV data collection. Neither automatic nor manual traffic counters are able to distinguish reliably commercial cars, vans, utilities or 4WDs. Conversely, both manual and the more sophisticated automatic counters can identify different sizes of medium and heavy trucks (typically over 3.5 tonnes gross laden weight).

It is not immediately clear what classification applies to the current WTSM.

The simplest approach for the update to WTSM appears to be to derive a model for all medium and heavy commercial vehicle trips. We would assume that the rest are covered in the commercial travel element of the person models.

Prior Matrix Accuracy

The existing WTSM matrix is of unknown reliability and its source is unclear. If we were to seek to improve the prior matrix, to avoid carrying forward biases into matrix estimation, there are two areas (at least) to consider:

- the distribution of trip ends: we might expect CV trips to be more focused in specific areas than car trips and the more discriminating our trip end estimates the better we might expect the trip matrix to be;

- trip lengths: given trip ends, the major determinant of the trip pattern is the average length of CV trips; independent verification of this would provide reassurance and could if necessary be used to develop an improved synthetic trip distribution.

Data Requirement

Arising from the above discussion, these are:

- the first main data requirement – an imperative - is an extensive set of screenline count sites identifying CV flows which can be used for matrix estimation;
- the second relates to trip ends; I see little point in attempting any general surveys, but consideration will need to be given to specific data collection for major generators (port, airport etc) – advice required; a classified cordon count around such generators could be used to provide an estimate of total in/out daily CV flows; given the cost this is probably only worth contemplating for generators with individually significant impacts on CV traffic in Wellington;
- the final data concerns trip length: while I can think of imaginative approaches, all would be costly to develop and pilot given the well-understood risks associated with most detailed CV survey methods; thus, such approaches would appear to be best deferred to later studies.

Growth Forecasting

It is probably appropriate to use the current trip end models, but it would be extremely helpful if evidence could be assembled on historic CV growth rates as a means of checking whether these simplistic models are indeed suitable. It would for example not be unexpected to find higher than expected growth rates on the motorways.

Future Developments

Substantial future developments would draw on any further studies by Transfund. Otherwise WRC would have to lead such developments (and this would be expensive).

4.13 Task 2.12 Use of Intercept Matrix Data

This is a scheduled for Task 12, but there is a need at an early stage to clarify and specify the strategy.

4.14 Task 2.13 Vehicle Types

We shall need to state at the outset how vehicles are to be classified and segmented and which are to be included in the personal travel model and which in the CV model.

4.15 Task 2.14 Model Structure Simplifications

This task is designed to determine whether the model can be simplified so as to exclude the need to develop 2 sub-models – the car ownership and family structure models. Most sub-models are based on person trips – the proposed production model (for most trip purposes), the attraction, distribution, mode choice and assignment models – and the planning data is based on population rather than households. The car ownership model is necessarily household-based. If the car ownership model could be foregone, and if family structure variables could be excluded from the trip end model, then all model inputs would be population-based, eliminating the need for a family structure model. This would much simplify the model and reduce the costs of setting it up.

The characteristics of a simplified model are outlined in Appendix 3.

4.16 Task 2.15 Park-&-Ride

The approach to modelling park-&-ride will be specified and agreed. Modelling sophistication is not required in this version of the model, although it may be developed further in future.

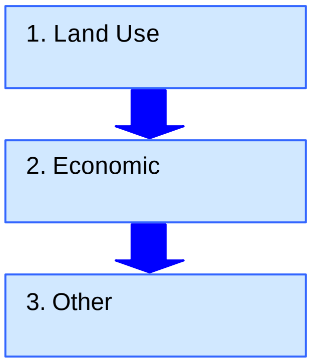
4.17 Task 2.16 Ports and Airports

Review how these are to be handled in the model, and the available external data (cf ART?).

5. Task 3 Inputs - Land Use, Economic and Other Data

5.1 Scope

To cover base and future years.



5.2 Task 3.1 Land Use

Inputs
MERA model.
Other (WRC-provided data)
New zone system (Task 4)

Processing
Specify required land use data – this will comprise not only demographic and employment data but other inputs such as school places and zone descriptors (parks, major zonal features, major shopping centres etc).

Review MERA Model. Ensure base is 2001.

Availability of forecasts of employment categories to be established.

Means of obtaining person/household types to be confirmed.

Process of developing base and future year data to new zone system to be finalised.

Subsequent tasks to be specified.

See Task 13.

Outputs
To new zone system:
- base land use data,
- process of forecasting future land use,
- future land use data;
- modelling procedures for inputting data.

Report.

Resources

5.3 Task 3.2 Economic

Inputs
External base values of economic factors.
External forecasts of economic factors.

Processing
Specify required economic factors. Eg fuel prices, public transport fares, incomes ....
Obtain data for both base and future years.
See Task 13.

Outputs
Economic values for base and future years.

Resources

5.4 Task 3.3 Other

Inputs
Specify.

Processing
Identify all other required inputs.
Assemble necessary data.
See Task 13.

Outputs
Other inputs for base and future years.
Note.

Resources

6. Task 4 Review, Amend and Implement Zone System

Inputs
Base & future planning data by zone and more detailed census subdivisions.
Base & future networks.
Zoning and network plans.

Processing
Check zone system compatibility with main data sources (census, other – *specify*) and with other, more detailed models
Identify large zones (by level of activity or area, in base or future years); specifically consider ‘sensitive’ areas:

- Wellington and sub-regional CBDs,
- developing areas,
- activity centres (shopping, recreation ...)
- access to central station,
- access to the rail network,
- future project corridors,
- access to important bottlenecks and the highway network, including representation of major turning movements.

Review network connectivity of these zones.
Decide whether any zones should be subdivided to improve assignment to the network or to reflect differential traffic growth patterns.
Design subdivisions – which should recognise the requirement to obtain planning data for any subdivision.
Consider renumbering zones so all ‘dummy zones’ for future disaggregation are grouped together rather than a series of ‘gaps’ in the real zone numbering. This would ease the identification of such zones and hence the ability to exclude zero-filled cells)
Include required network changes in Task 5.
Produce base & forecast zonal planning data files to new zone system (in Task 3).
Design sector aggregations for use in model calibration and validation (eg 5, 10 and 20 sectors).
See Task 13.

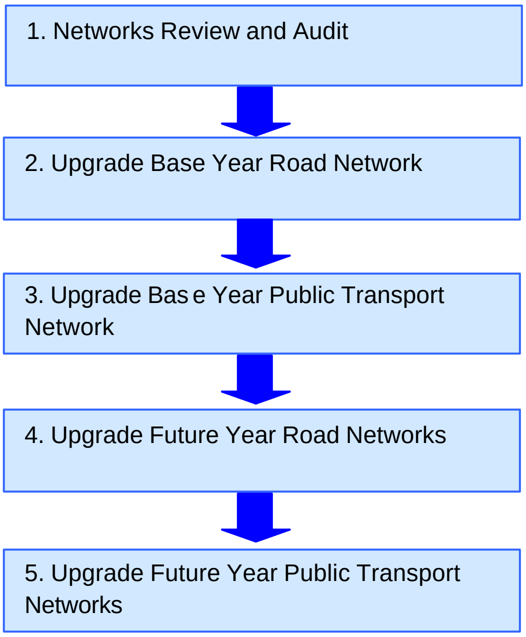
Code all surveys to the new zone system – *specify process.*

Outputs
Revised zone plan.
Revised zone numbering system.
Technical Note.

Resources

7. Task 5 Networks – Public and Private

7.1 Scope



7.2 Task 5.1 Networks Review and Audit

Inputs
Current networks
Revised zone system
Network inventory data; GIS plans etc

Processing
Major review of:
- coding protocols (emphasised to be important)
- source of and accuracy of existing network data (from GIS/TRACKS)
- network density (presently though to be ok)
- centroid connectors in CBDs (important)
- new zone system

For road:
- capacities
- speed/flow curves
- representation of key road bottlenecks (a few key locations envisaged)
- network expansion at key intersections (as above, for bottlenecks only; no major intersection modelling)
- car parks will not be handled in networks, but in generalised cost
- walk links for 1-way streets in CBD

For public transport:
- service specifications
- running times
- interchange representation
- access to all stations, park-&-ride (analysis of household survey and rail survey?); very important regarding future park-&-ride provision
- fares

Outputs
Plan for network upgrading.

Resources

7.3 Task 5.2 Upgrade Base Year Road Networks

Inputs
As 2.1.

Processing
- if necessary, overplot network on GIS plans to check structure
- if necessary, compare coded link distances with crow-fly calculation on GIS node coordinates?
- compare junction and link network geometry with survey data and junction sketchplans?
- if necessary, merge network data files for different time periods (& years): apply consistency checks.
- correct junction and link data files.
- review attributed capacities; revise on experie nce (or survey) basis.
- allocation of link-specific variables rather than the use of global values or values hard-coded in CDFs (e.g. as free-speeds, lane capacities, link-specific routing parameters, link tolls
- review speed/flow curves; revise as above.
- extend network to connect with revised zoning system
- test/validate the networks using test paths etc.
- verify & update slow mode networks
- development of simplified junction-modelling procedures is rejected
- develop a data-storage system with relevant file naming, version numbering, data extraction macros, directory structure and QA procedures

Outputs
An inventory library of all significant junctions.
A common database of all base and future DM networks.
Revised model network data files for base year.
A simplistic junction modelling system, including network coding
A data storage and management system
Technical note.

Resources

[11] There are significant advantages in maintaining a link-only assignment model without using turn-penalty functions.

7.4 Task 5.3 Upgrade Base Year Public Transport Networks

The BAH documentation indicates that the PT network was recently updated to 2000 and hence a major upgrade is not anticipated. However, there are issues with version identification and definition of period of the year which the networks should represent.

Inputs
- Timetables
- GIS overlays
- Existing model networks

Processing
- overplot network on GIS plans to check structure?
- if necessary, compare coded link distances with crow-fly calculation
- review transit speed/flow curves, and interface with auto-assignment
- extend network to connect with revised zoning system
- review mode and vehicle type definitions against current an potential segmentation (e.g. express, all-stops, flyer buses, HOVs etc)
- review approach to park & ride assignment (recognising that the approach may be developed further in later model versions)
- test/validate the networks using test paths etc.
- verify modelled journey times against timetables, ETM data or observation, particularly in CBD

Outputs
2001 Network.
Technical Note.

Resources

7.5 Task 5.4 Future Year Road Networks

Inputs

- RLTS
- Committed projects
- State Highway Strategies?
- Assumed upgrades in current model

Processing

The general approach is to include expected future upgrades in an agreed sequence as well as minor upgrades required to obtain feasible, stable assignments

- Identify assumed future upgrades in current model (and verify against documentation)
- 'Workshop' all known/possible projects with all RCA's to get an agreed staging of projects (including future parking charges (and capacities?)). As well as general capacity upgrades, the workshop should also cover projects such as possible HOV facilities, bus priority measures, toll/charging locations etc.
- Specify coding protocols
- Develop coding for network upgrades – use network modification system so that individual projects can be 'mixed & matched' as required
- Incrementally test each forecast year on preceding years network (i.e. assign 2006 demand to 2001 network) – identify network deficiencies (large delays, low speeds, large increases in cells of skimmed time matrices)
- Identify if any known project is likely to solve the deficiency and compare with agreed sequencing from workshop
- If no known project exists, develop feasible solution (general rule of thumb is that such 'unknown' upgrades should be restricted to intersection upgrades with link upgrades only being used if previously identified as future projects)
- Test upgrades and finalise network for forecast year
- Repeat for subsequent forecast years

Outputs

- Library of future network improvements
- Agreed sequencing of major projects
- Documentation and coding of minor network improvements for each forecast year to obtain stable, feasible models
- Documentation

Resources

7.6 Task 5.5 Future Year Public Transport Networks

Inputs
RLTS
Assumed upgrades in current model

Processing
General approach is only include committed upgrades (any??). The PT modes have little or no capacity restraint (except where bus speeds are influenced by congested car speeds), and hence future upgrades generally not required to obtain feasible and stable assignments.

- Identify assumed future upgrades in current model (and verify against documentation)
- 'Workshop' all known projects with all RCA's to get an agreed staging of projects (including future fares)
- Develop coding for network upgrades – use network modification system so that individual projects can be 'mixed & matched' as required
- Review bus speeds for each forecast year
- Review changes in mode split driven by relative change in car/PT costs

Outputs
- Library of future PT network improvements
- Agreed sequencing of PT upgrades
- Documentation

Resources

8. Task 6 Car ownership

8.1 Key Issues

The following discusses a number of key issues.

8.1.1 Cross-sectional and temporal models

The current WTSM car ownership model is a cross-sectional model estimated on a single household interview survey data set. Such models have been widely used, and are effective in reflecting the variation within the data in terms of variables such as income, household structure and location. The current model is estimated separately for seven household categories, which are amalgamations of the 18 categories used in the trip production model. Within these categories, account is taken of income and “accessibility” variation.

As is conventional within such structures, the model is developed as a sequence of conditional probabilities. A first model splits the households between car-owning and non-car-owning, and a second then splits the car-owning households into single and multi-car households.

General experience is that while the cross-sectional models retain the ability over time to reflect the **distribution** of car ownership (and specifically the variation between zones and household types), they are not usually successful in representing changes in the **level** of car ownership. Put another way, the cross-sectional model is unable to allow for trend effects which are unrelated to the above-mentioned cross-sectional variables, and it is therefore necessary to import some corrective trends.

Ideally, this would be done using a cross-sectional data set which contained data for different years, possibly of a panel form, but more likely as a “repeated cross-section”. For example, in the UK, where the Family Expenditure Survey [FES] is carried out on a continuous basis, it is possible to estimate car ownership models using cross-sectional data over a period of nearly 30 years. Analysis of this kind has been able better to determine the sort of trends which are required in addition to the purely cross-sectional relationships.

In the case of Wellington, some information along these lines might be obtained by pooling the data from the 1988 and 2001 household surveys. However, clearly a reasonable number of points are necessary to establish a convincing trend. In addition, the UK FES has the advantage of consistent definitions over time: experience shows that even with the best of intentions, surveys repeated at long intervals tend to have problems in this respect. A more promising Wellington alternative would be to obtain current and past data from the 5 year censuses, if mutually consistent and available at reasonable cost.

In the absence of suitable repeated cross-section data, it is necessary to apply more coarse trends on the basis of aggregate time series data relating to the growth of car ownership.

Clearly, these “trends” may be partly modellable. For example, we can expect an effect from car price movements (see below), and from the level of driving licence holding. There is a preference towards reflecting such variables explicitly, as far as possible, not least because this can improve the ability to model policy. It must be recognised, however, that the data may not be adequate for this.

The main point to stress is that it is not reasonable to expect a cross-sectional model to forecast correctly without adjustment.

Car prices
Since car prices, either by policy or because of macro-economic conditions, may move differently from the general experience of the past, it is certainly useful if the car ownership model can respond explicitly to this variable.

For obvious reasons, there is no possibility of estimating the effect of car prices from cross-sectional data. Moreover, it has proved quite difficult to estimate the effect from time-series data, given the correlation with other variables. However, some successful studies have been carried out, and there are also methods relating to vehicle type which can make use of the greater variability of prices relating to different kinds of vehicle.

Given a credible elasticity, the best approach is to incorporate the price term in the model in a suitable form, and calibrate a coefficient which delivers the accepted elasticity. This has been done in the UK, using a modelling structure essentially similar to that in the current WRC model.

As implied earlier, it is preferable to allocate as many of the temporal effects as possible to explicit factors (such as car prices). The residual change over time can then be represented by an unexplained trend.

Licence-Holding
An additional effect of increased licence-holding on the trend in car ownership growth has been specifically incorporated in some model systems. Explicitly or implicitly these models represent how licence holding is expected to increase through a population cohort effect: the current generation of younger people is far more likely to have licences than previous generations, and will retain these licences as they age, thus increasing licence holding rates in older population cohorts in the future.

This effect may be most marked for the retired household segments which are to be specifically modelled, and may be expected to influence the future rate of car ownership growth will be specifically forecast. This will need to be considered in the process of adjusting the model constants in future forecasting.

The BAH 1997 Trend Forecasts and Model
A respectable review of car ownership trends in NZ and the development of a temporal forecasting model was done by BAH for Transit NZ in 1997. This model also included a car price term. For efficiency, before we consider anything more ambitious, this model will be reviewed against international practice and the most recent trends (which have also been affected by car price changes). It may be a suitable basis for temporal controls on the car ownership forecasts, perhaps with some minor updating.

8.1.2 3+ Car Owning Households

The rising levels of car ownership in some household types brings with it the possible need to distinguish between households with 2 cars and those with 3 or more. In principle, this can be done by extending the conditional approach in a straightforward manner. In respect of the current model, an attempt was made to do this, but there was insufficient data to produce significant coefficients. It will need to be assessed whether this conclusion will still apply to the new survey data.

Since in any case it is not expected that households with 3+ cars will be a major segment, it will be useful firstly to determine the sensitivity of the overall model to such a distinction. Crucially, it will depend on the difference in trip productions between households of the same classification but varying between 2 and 3+ cars.

Aside from these essentially practical questions, however, the modelling does not present any significant problems, and is effectively an extension of the existing specification.

8.1.3 Accessibility terms

It is well-known that, other things being equal, car ownership levels are lower in urban areas than in rural. A major study in the UK (RHTM, 1978) established that there was a more or less continuous decline with measures of urbanisation, such as net residential density. A somewhat better relationship was found using a measure of “accessibility”.

It was not clear, however, whether such measures represented a convenient summary measure of urbanisation or whether they had some explanatory power *per se*. For example, if residential density declined in a given area, without any other changes, would we expect an increase in car ownership? The approach taken in the UK was merely to use the measures to **classify** zones into broad bands – the measures themselves were not recalculated as a result of changes over time.

One particular reason for this stems from a naïve interpretation of the accessibility measure. If, for example, we define “accessibility” as a standard “Hansen” index : $A_i = \Sigma_j E_j \exp(-\lambda G_{ij})$ where G is car generalised cost, then we expect a declining propensity to car ownership with increasing A. However, increasing congestion in urban areas will lead to a decline in A, suggesting (counter-intuitively), that car ownership will **increase**.

The development of more explicit theories of car ownership has introduced the concept of “differential accessibility”. In other words, part of the justification for owning a car is in terms of the additional accessibility which may be obtained, over and above that afforded by non-car modes (including public transport). If this is carefully specified, then it avoids the counter-intuitive predictions just mentioned.

The treatment of accessibility in the current models is partly motivated by such arguments, but a complete rationale is lacking. At the least, it needs to be demonstrated that sensible outcomes will be predicted in respect of straightforward policies such as higher fares or fuel tax. It is recommended that further thought be given to this, and this may impact on the current application, in which accessibilities are recalculated for each future year but not in respect of particular “policies”, so that they remain outside the main iterative supply and demand loop. The following represents a preferred approach. It replaces the two accessibility measures in the present model by a single differential accessibility based on the same model variables (employment and generalised cost by mode).

Consider that there are three modes: 1 = walk/cycle, 2 = public transport, 3 = car. For convenience, we assume a model of mode choice conditional on destination. For someone <u>confined</u> to the "slow" mode, the utility of a journey from i to j can be represented by the sum of the utility of the destination U_j and the (dis)utility of the travel from i to j by mode 1 U_{ij1}. In the usual way, we assume a generalised cost approach with

$$U_{ij1} = - ?\ GC_{ij1}$$

Consistently with general theory, we will assume that the dominant "attraction" of a zone is the number of people employed, though this could be changed without difficulty. We make use of the theory of "size" variables, whereby $U_j = \ln(E_j)$.

To represent a possible hierarchical structure between destination and mode, we introduce a structural parameter θ, where $0 \le \theta \le 1$. Hence we write:

$$TU^{[1]}_{ij} = \ln(E_j) - \theta\ ?\ GC_{ij1}$$

where TU represents the "total utility" of the journey, and the [1] indicates that only the walk mode is available.

For someone who could make use of either walk or public transport, we need to substitute the appropriate <u>composite</u> generalised cost in this formula. For convenience we will assume a single level mode choice model, but a more complicated structure could be catered for (once the number of modes exceeds 2, that is). Thus we use the formula:

$$TU^{[12]}_{ij} = \ln(E_j) - \theta\ ?\ GC_{ij*[12]}$$

where the composite GC is defined in the usual way as:

$$GC_{ij*[12]} = -1/?\ \ln [Ó_{m=1,2} \exp(-\ ?\ GC_{ijm})]$$

Finally, consider someone who, in addition to slow modes and public transport, also has a car available. In an exactly comparable way we use the formula:

$$TU^{[123]}_{ij} = \ln(E_j) - \theta\ ?\ GC_{ij*[123]}$$

where the composite GC is defined as:

$$GC_{ij*[123]} = -1/?\ \ln [Ó_{m=1,3} \exp(-\ ?\ GC_{ijm})]$$

As usual, it is not possible for the Generalised Cost to <u>increase</u> as a result of more modes being available. The extent to which it decreases depends on both the value of ? and the relative performance of the added mode.

So far we have focussed on a single destination. We now need to extend the process, effectively by compositing over all possible destinations available from a given origin. This again follows standard procedures, and we use the formula:

$$TU^{[K]}_{i*} = \ln [S_j \ \exp (TU^{[K]}_{ij})]$$

where K represents the set of available modes.

The additional utility gained from car ownership for residents at origin i can then be written as:

$$\Delta U[car]_i = TU^{[123]}_{i*} - TU^{[12]}_{i*}$$

This can be converted into GC-like units by dividing by ($-\theta$?).

This provides a principled way of introducing the accessibility effect into the car ownership model. It has many of the features of the existing model, but it represents a theoretical improvement, while the effort required in calculation is essentially similar.

8.1.4 Family and work structures

It may be noted that although the current model estimates separate models for each of the grouped household classes, modelling work in the UK has suggested that the impact of household structure tends to be confined to a single variable (either the income coefficient or the constant term). This means that the models for all household groups can be estimated simultaneously on the pooled data set, and also facilitates the testing of significant differences between different household categories.

8.1.5 Fitting to census data

While cross-sectional models of the WTSM form capture the general variation within the sample, it has often been found that they perform less well at the zonal level (in spite of the earlier remarks about accessibility etc.). This suggests that there may be specific features of the zones (for example, the availability of on-street parking) which are not captured in the model variables.

In fact, the application of the current model at the zonal level was judged to be acceptable. Nevertheless, it would be prudent to allow for a correction, based on aggregate car ownership figures for each zone from Census data. There are various ways in which this correction can be made, depending on the extent of available data. In the past, given that zonal income data was not usually available, the model was often calibrated to the zone by adjusting the mean income to fit the base year car ownership. Alternatively, the model constants can be adjusted to obtain the same effect.

8.1.6 Implementation

Although the model is calibrated at a household level, it needs to be applied at the zonal level. Most of the questions of implementation relate to this.

Whatever household segmentation is adopted for the model, it will be necessary to produce zonal forecasts of the **numbers** of each household segment. This process is described in more detail in Task 7.6.

Given the numbers of households, the essential task is to allow for the **distribution** of household income. The "traditional" way has been to assume a standard mathematical form for the income distribution (eg gamma or lognormal), and given a mean zonal income (which was usually calibrated to reproduce base year zonal car ownership), integrate the car ownership function against the income distribution.

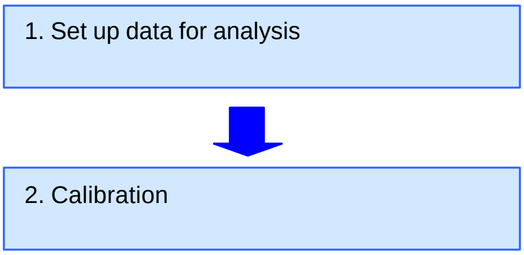
However, while distributions such as gamma may give a passable representation of the income distribution of all households taken together, they are much less well suited to specific household types. An attractive alternative is sample enumeration. Given a household sample of the relevant category with their actual incomes, we can apply the car ownership model directly. Income growth can be allowed for at the individual household level. We then take the average values of p0,1,2,3+ for all the sampled households.

This approach is applied separately for each zone. The model will have been slaved to reproduce 2001 census car ownership, so there are some different model parameters for each zone. These zone-specific models are run, separately for each household type, using a single city-wide income growth.

Note that because car ownership proportions are produced separately for each household type, and the household types are themselves defined in terms of different person types, this automatically produces the forecasts required by the trip production model of person type by household car ownership (and possible other household characteristics)

8.1.7 Scope

The tasks are:



8.2 Task 6.1 Set up data for analysis

Inputs
Household survey
Census data for 2001 (zonal cars and income)
Historic data: census and car price data
WTSM network models (query: peak, interpeak or both?)

Processing
Household survey files for disaggregate calibration.

Census data for base year fitting.

Review BAH/Transit report/model on car ownership trends; determine what additional data is required for historic data for trend analysis and trend parameters.

Network models for accessibility terms (initial calibration based on old WTSM model (?) subsequently updated when new sub-models have been developed) – construct differential accessibility measure.

Outputs
Data files.

8.3 Task 6.2 Calibration

Inputs
Data files.
LIMDEP calibration software

Processing
The model concept is illustrated in Figure 8.1. The mathematical structure is the same at leach level in the sequential model. Take for example the choice between 1 and 2+ car ownership:

$$P_{2+|1+} = S_{2+}/[1 + \mathrm{Exp}\,(LP)]$$

where:

$P_{2+|1+}$ is the probability of owning 2 or more cars for the group of households owning at least 1 car;
S_{2+} is the saturation level of this probability (≤ 1)
and LP is called the linear predictor.

and:

$$LP = \alpha.f(I) + \beta.Acc_z + \gamma + \lambda_z + \delta$$

where:

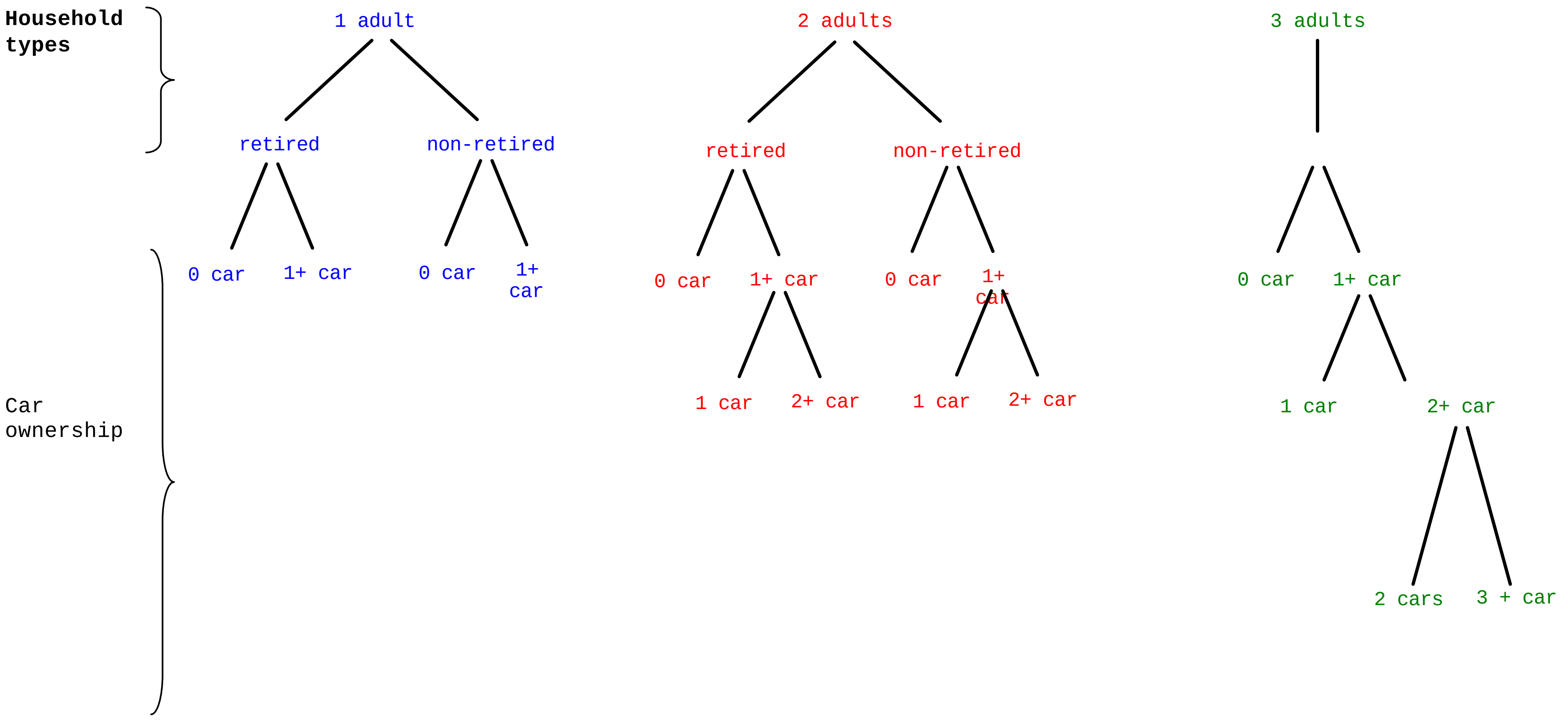
α is the coefficient of some function of household income I (either income, log income or square root of income – established through statistical analysis)
β is the coefficient of differential accessibility Acc_z for zone z
γ_y is a temporal trend adjustment for each year (eg related to car price and licence-holding, or the BAH/Transit model)
λ_z is an adjustment to fit the model to zonal census car ownership
δ is a constant.

The calibration tasks are:

(i) through tabulation and preliminary statistical analysis verify model structure (see also Task ...)
(ii) calibrate models vs income, and test for best income variable and “saturation levels”: [0/1+], [1/2+],?[2/3+]; use LIMDEP
(iii) test inclusion of differential accessibility;
(iv) decide on modellable contributions to trend, and appropriate model form for inclusion: calibrate to specified elasticities; it is intended that this should be done in relation to the BAH/Transit model;
(v) calibrate base year car ownership to zonal census data;
(vi) See Task 13

Outputs
Car ownership model specification.
Report.

■ **Figure 8-1 Proposed Car Ownership Model Structure**



9. Task 7 Trip Ends

9.1 Trip End Model Specification

9.1.1 General

Trip purposes:	segmentation (see 2.2.1)
Modes	all modes combined
Time period	24 hours travel

Note that the travel of children aged less than 5 years is excluded.

9.1.2 Trip Productions

The following characteristics which explain variations in trip-making are common to most recent trip generation models:

- person characteristics: child, unemployed adult, employed adult, retired person (see Chapter 2 for the proposed 7 person types);
- household characteristics: person type composition, number of cars[12] and household location (eg city, other urban, rural): *need to specify categories.*

Additionally, we are asked to look into:

- work structures (essentially distinguishing those who do not commute regularly).

While all but the last factor are represented in the present model, the household-based generation categories are generally ambiguous regarding the number of persons in the household, which seems unsatisfactory.

Our expected model specifications are very broadly as follows.
HBW ~ no. of employed adults
HBEd ~ no. of schoolchildren + no. of young adults
HBSh ~ all person types, household influence
HBSo ~ all person types
NHBO ~ all person types, employed persons
BU ~ employed persons

For all segments the number of household cars is a relevant factor for defining captive/choice (Chapter 2).

[12] In some models, also the number of drivers.

9.1.3 Scope

Trip Productions:

1. Create analysis files

↓

2. Statistical Analysis

Trip Attractions:

3. Create analysis files

↓

4. Statistical analysis

↓

5. External Trip Ends

↓

6. Family Structure Model

9.2 Task 7.1 Create Trip Production Analysis Files

Inputs
Household survey.

Processing
Specify required file contents for trip production analyses. Data should be all residents trips whether internal or external to the study area.

Create a flat file from the household survey, 1 record per person, including the necessary household data; for example:

- all-mode person trips per day by modelled trip purpose for a weekday, Saturday & Sunday
- person characteristics:
  - employment status (including full/part time, work arrangements and type of employment)
  - age
  - other individual characteristics if appropriate
- family structure characteristics:
  - family size
  - income
  - car ownership
  - other family characteristics if appropriate
- location (eg urban/rural ....)
- other, as appropriate.

Outputs
Documented household & person flat files.

Resources

9.3 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. For reasons discussed below the present model is unsatisfactory and, whatever specification is preferred, should be replaced. 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.

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 models.

The major complication potentially lies with the family structure model which needs to relate population based planning data to car ownership and family structure – but this complication applies more-or-less to any model specification. It may be reduced if it can be handled within the MERA software.

Inputs
Estimation files.

Processing
Choose statistical software: SYSTAT or LIMDEP.
The model specifications below are non-standard but draw on other mode 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.

HBW model:

Variations in trip rates are likely to be as follows:
- person effects: expect to vary by work structure (full/part-time; contractor ...[13]) because these dictate need to make a commuting journey;
- household effects: number of cars possible but unlikely; inverse correlation with number of children (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. However, although the effects of household car ownership are likely to be minor, it may be worth retaining this segmentation for consistency, given that the segments need to be differentiated for the mode choice.

Model form:

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

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

  Work trips/zone = Σ_{at} [NF_{at}*WF*α_{at} + NP_{at}*WP*β_{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

  α_{ae} β_{ae} 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 Chapter 2. 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).

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 NZ schooling regulations, and/or by tabulating schooling probabilities by age group using the household survey, census or education statistics; issue of pre-school and kindy;
- we may need to split these into choice/competition/captive.

HBSh and 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[14];
  - 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 =

  $$\Sigma_i N_i * T_i + \Sigma j(N_0 * \Delta T_0 + N_1 * \Delta T_1 + N_2 * \Delta T_2 + N_{3+} * \Delta T_{3+})$$

  where:

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

  $N_{0/1/2/3}$+ are the number of persons in the zone in households of car ownership level 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 (correlations between the actual and incremental trip rates, the model would become significantly more complicated).

To obtain the required split of trips between captive, competition and choice, the 4 person types need to be split for each zone into these categories.

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.

[14] Income is an interesting variable but with high correlation with car ownership, person type and family size and posing additional forecasting complications. For these reasons it is not included in the specification.

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).

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[15];
- produce range of model calibration fit statistics (eg R^2, t-stats, predicted vs observed at zonal level and sector aggregations);
- report.

Outputs
Trip production model.
Report.

Resources

[15] The best models of this type do not reproduce many of the geographic variations in the data. It is important that these are identified and checked and factors established to correct for significant differences (although we may not use the factors).

9.4 Task 7.3 Create Trip Attraction Analysis Files

Inputs
Household survey.
External survey (road).
PT surveys.
School survey
Land use data for 2001.

Processing
Specify required file contents for trip attraction analyses.

Create a flat file, 1 record per zone including:

- residents internal all-mode trips attractions by purpose from the household survey by purpose;
- non-residents internal attractions by purpose from the external road survey and public transport survey;
- zonal characteristics from the land use data;
- add any other zonal information (zone types, type of retail destination, presence of major recreational attractors, urban/rural etc); see Task 3.1.

Outputs
Documented flat file.

Resources

9.5 Task 7.4 Trip Attraction Statistical Analysis

Inputs
Estimation file.

Processing
Software: SYSTAT.

Models will be zonal attractions expressed as a linear function of planning variables plus a contribution from zone types either as a segmentation or as additional variables. Typically the variables will be:

- HBW: employment
- HBED: schoolplaces plus?
- HBSh: retail employment, other employment and population?
- HBSo: as HBSh?
- NHBO (as O/Ds): almost anything
- BU: employment by type
- other special zonal characteristics (important – Malcolm Douglas).

Calibration:

- it may be sensible to use aggregate geographical areas rather than zones, perhaps 30 or 60 areas (review)
- the calibration process is regression of observed attractions against planning data, but this is a tortuous process:
  - the statistical measures are unreliable and biased for this sort of data, so some care should be exercised in choosing the explanatory variables;
  - it is always worth graphing the relationships before accepting the models;
  - it is always worth looking at residuals and removing outliers before firming up on a model; significant outliers should be specifically studied for special zonal characteristics;
  - if observed values cover a very wide range, it is often useful to estimate models for sub-ranges (eg <10,000 trips, >10,000 trips);
  - produce correlation matrices: be very wary about high correlations between the explanatory variables (which means that the statistics may not be able to distinguish between 2 or 3 explanatory variables) so it must rest on the analyst's judgement;
- check that the estimated total attractions are in an appropriate relationship to the productions;
- analyse models for geographical bias/fit based on some predefined geographic aggregations and/or the sector aggregations; incorporate any geographic k-factors;
- produce range of model calibration fit statistics (eg R^2, t-stats, predicted vs observed at zonal level and sector aggregations);
- report.

Outputs
Trip attraction models.
Report.

Resources

9.6 Task 7.5 External Trip Processing and Forecasting

Inputs

Processing
To be determined. Essentially growth rates will be based on factors such as population, employment, economic growth and car ownership. Ideally, some historic data would be obtained to provide evidence of growth rates. Possible default would be to use the overall study area travel growth.

Outputs

Resources

9.7 Task 7.6 Create Family Structure Model

9.7.1 Concepts

The car ownership and trip end models require disaggregated family structure distributions for forecasting (a process which many NZ models seek to short-cut). The requirements are as follows.

Person Types
The basis of the model is the population split into 7 person types which we expect MERA will forecast for each zone. For convenience (as we shall see) we have split up category 7 to identify all workers.

Person types
P1: infant (<5) P2: child (5-16?) P3: young adult (17-25) unemployed P4: young adult (17-25) employed by: P5: adult (26-60) unemployed P6: adult (26-60) employed P7a: adult (>60) unemployed P7b: adult (>60) employed

Household Types
The car ownership model uses a split of households by type, by zone.

Household types
H1: retired, 1 adult H2: non-retired, 1 adult H3: retired, 2 adults H4: non-retired, 2 adults H5: 3+ adults

For the trip production model, and to allocate the car ownership levels, we need the classification of population by household and person type (although the same distribution could be applied for each zone).

Population distribution	Person Types							
Household types	P1	P2	P3	P4	P5	P6	P7a	P7b
H1								
H2								
H3								
H4								
H5								

We will review the existing MERA models, to see to what extent it is reasonable for these to produce the required forecasts. However, we present in 9.7.2 an alternative methodology which could be developed.

Income
For forecasting car ownership, we need to take account of the income distribution by household type for each zone. We propose that this will be achieved by sample enumeration (see 8.1.6).

Car Ownership
For the trip end model we need a classification of household types by car ownership for the captive/choice segmentation. This will be estimated by the car ownership model.

Household Types		Car Availability
H1.1	0 cars	Captive
H1.2	1 car	Choice
H2.1	0 cars	Captive
H2.2	1 car	Choice
H3.1	0 cars	Captive
H3.2	1 car	Competition
H3.3	2+ cars	Choice
H4.1	0 cars	Captive
H4.2	1 car	Competition
H4.3	2+ cars	Choice
H5.1	0 cars	Captive
H5.2	1 car	Competition
H5.3	2 cars	Competition
H5.4	3+ cars	Choice

For the trip end model, for each zone we further need a cross-classification of persons within household type, the latter segmented by car ownership level. This is obtained directly by combining the household car ownership forecasts with the Person type by Household type distribution.

Households by Type and Car ownership	Person Types P1	P2	P3	P4	P5	P6	P7a	P7b
H1.1								
H1.2								
H2.1								
H2.2								
H3.1								
H3.2								
H3.3								
H4.1								
H4.2								
H4.3								
H5.1								
H5.2								
H5.3								
H5.4								

Again this will be based on sample enumeration based on the above row and column totals or within the MERA person/household categories (see 9.7.2).

Employment Type etc
For the HBW trip production model the workers are segmented in relation to full/part time employment, employment type and work arrangement. It is proposed to assume that these proportions do not vary by household type.

Employment type is unlikely to be forecastable and will therefore be a scenario variable unless evidence is available on trends.

Other
We may also be interested in:
- households by size;
- scholar/student propensities (by age group).

Analysis
Prior to defining the analysis the details of the approach need to be finalised in discussion with MERA.

For the sample remuneration and any other specific model disaggregations, the source of data is the household survey.

This structure serves both the weekday and weekend models.

9.7.2 An alternative approach for forecasting the number of households by type for each zone

The method relies on a base distribution, or "template", of different household types in the zone (or over the whole study area) which are then re-weighted with the aim of achieving certain "targets".

The aim of the procedure is to produce separately for each zone a set of numbers of households in each type required by the car ownership model, in a way which departs as little as possible from the current zonal distribution by household type, but which reflects the changed forecasts of population in various categories. In making the adjustments, the procedure takes account of the distribution of person types within each household type, as set out earlier.

As a result of this, each zone is given forecasts of the number of households of each type in such a way that the "target" population is met (to an acceptable extent). Because we know the person type composition of each household type, we have the required person x household distribution at the zonal level. In other words, we know the number of persons of each type in each household type.

If a sample enumeration process is being used for car ownership, then the ratio of the required number of households of a given type to the number of such households in the sample produces the weight for each sample household. Taken all together, therefore, the estimated household totals are reproduced, segmented by levels of car ownership. The car ownership levels for the various household types can be combined to the groupings required by the trip production model.

Taken together, we obtain the required cross-classification of peron type by household car ownership group. At the same time, other aspects of the household to which the person belongs (for example, the number of adults) is also available. Thus this process produces precisely what is required for the trip production models

We will give this method and the possibilities of the MERA based forecasts careful consideration, bearing in mind that neither car ownership nor household/person type changes are likely to be particularly significant in future. Car ownership is already high and the forecasts are for no great changes in family size. This may predispose us towards technical compromise. Given the cost trade-offs here, we must be sure that there is no much cheaper way of doing something where the technical losses are acceptable.

10. Task 8 Distribution and Mode Choice

10.1 Task 8.1 Create Analysis Files/Set up Statistical Procedures

Inputs
Survey Data Base.
LIMDEP

Processing
Specify requirements: the calibration process requires a host of files of varied structures, a wide range of diagnostic outputs and a complex sequence of model estimations using LIMDEP. Eg:

Files:
- generalised cost matrices by mode and time period;
- trip matrices and trip ends by mode, time period and segment.

Diagnostics:
- sector trip end and trip matrix aggregations;
- trip length/cost distributions;
- statistical tests;
- geographical aggregations for model calibrations diagnostics.

Estimation:
- LIMDEP set-ups to be established, and diagnostic outputs;
- sequence of estimation steps and decision points to be specified.

Outputs
Files.
Output diagnostic procedures.

Resources

10.2 Task 8.2 Develop Model Specification

Inputs
Household survey data files.
Generalised costs by mode and time period.
LIMDEP (?).

Processing
The model specification options have to be constrained in order that this task is kept within manageable limits. The anticipated specification is as follows in the table, but first we introduce the key issues:

- this is unquestionably by far and away the most demanding pair of models to calibrate; there are also a large number of possible specifications which can, if not controlled, cause budgets to be over-run; they are also demanding on calibration software;
- it is necessary to recognise and anticipate the choice between pre- and post-distribution mode choice in the work programme; it is also necessary to establish a potential specification for each;
- the approach to time periods needs to be considered; we favour combining the data for a 24 hour model and using matrix factors to generate the time period matrices for assignment; for efficiency, it is sensible to input peak period network data for those purposes occurring primarily in the peaks (eg commuting etc) and interpeak network data otherwise although this involves some compromise in that not all purposes split conveniently into one or other time period; it may be feasible to consider combining peak and off-peak costs in proportions varying by trip purpose;
- generally, the complex estimation process needs to be specified and set-up;
- there are a lot of detailed model specification issues to consider, including:
  - how external trips & non-residents are handled;
  - whether all trip purposes deserve this complex model hierarchy;
  - how to deal with difficult modes: car passengers and slow modes;
  - issues relating to costs: intrazonal costs and parking costs;
  - specification of composite cost links for non-generic model structures if required;
- because the data is sparse, the level of aggregation and segmentation of the estimation process needs careful design;
- referring to Task 8.1, the number of files required is substantial and the automated pre-assembly of the data is therefore a major and critical task.

Outputs
Preferred Specification.

Resources

General Specification Issues

Issue	Approach	Comments
Structure	See Figure 10.1	This structure is that used in the London model and may have the attributes needed for Wellington. It will be necessary to confirm its validity as against a post-distribution structure (not illustrated, the present model).
Time period	24 hours	Demand models built on a 24 hour basis, with the matrices later factored to the individual time periods. This reduces number of models to be calibrated and avoids estimation on small data sets (eg interpeak HBW).
Networks[16]	AM peak: HBW & HBEd Interpeak: all other segments	Consider feasibility of combining costs for each time period in purpose-specific proportions
Estimation	Coefficients on unexpanded data. MSCs and equivalents on expanded data.	
Statistics of estimation	Specify requirements	Need to confirm the software to be used for estimation.
Single/double constraint	Suggest: ❑ HBW & HBEd doubly-constrained[17] ❑ other purposes production-constrained	While it is clear that work and education trips are constrained to the number of work and school/student places, no such attraction constraints exists for other purposes.
Detailed model specific-ation issues:	❑ model all resident and non-resident's trips (ie including externals) ❑ are the non home based and business trip purposes best done by Furness growth factor techniques being less sensitive to network issues? ❑ pre- or post-distribution mode choice? ❑ car passengers with car drivers or separate mode? should it be a sub-model choice model or car occupancy factor? ❑ how slow mode trips are handled (being mainly intrazonal) in the sub-mode choice model ❑ parking costs ❑ geographic MSCs – identification of an appropriate geographical structure ❑ intra-zonal trips/costs	

[16] We had hoped to treat captive as a single segment, but this seems difficult when it includes a mixture of peak and interpeak trips.

[17] For doubly-constrained models only, attractions are balanced to total productions using an overall factor; for singly-constrained models, in forecasting the attractions would be treated as 'attraction factors'.

Pre-Distribution Mode Choice (Figure 10.1)

Issue	Approach	Comments
Production Mode split	Input Productions: by segment ❑ HBWa: Ca, Co, Ch ❑ HBWo: Ca, Co, Ch ❑ HBEd: Ca, Co, Ch ❑ HBSh: Ca, Co, Ch ❑ HBSo: Ca, Co, Ch ❑ NHBO: Ca, Co, Ch ❑ BU: Ca, Co, Ch Modes: at this level, a 2- mode model: (1) car passenger & driver combined and (2) public transport and slow combined. Composite costs for each of the 2 modes are common to each captivity segment and extracted from the distribution model.	Extent of captive/choice segmentation depends on pre-analysis. Zonal data is likely to be too sparse for reliable estimation given low proportion of PT trips, so: ❑ either aggregate zones on some basis (eg by location, or by differential accessibility) ❑ or calibrate at a disaggregate level[18] updating the ASCs on aggregate data, ❑ and combine captivity sub-segments in calibration, perhaps only varying ASCs for sub-segments.
Distribution by mode	Segments: purpose (7) and mode (2) Input productions by mode (aggregated over car availability). Input attractions by purpose [modes compete for zonal trip attractions]. Composite costs for each mode by matrix cell are derived from the sub-model choice models for each purpose.	With fewer segments and model parameters, zonal distribution calibration should pose few problems.
Sub-Mode Choice	To be specified.	

Post-Distribution Mode Choice (Figure 10.2)

Issue	Approach	Comments
Distribution	Input Productions: by segment ❑ HBWa: Ca, Co, Ch ❑ HBWo: Ca, Co, Ch ❑ HBEd: Ca, Co, Ch ❑ HBSh: Ca, Co, Ch ❑ HBSo: Ca, Co, Ch ❑ NHBO: Ca, Co, Ch ❑ BU: Ca, Co, Ch Input Attractions by purpose [car availability segments compete for attractions]. Composite costs are extracted from the mode choice model separately for each segment.	Extent of captive/choice segmentation depends on pre-analysis. Some concern about the sparseness of the captive and competition sub-matrices being insufficient for deterrence function determination. Not sure if there is any solution to this apart from amalgamating these segments.
Mode Choice	Segments: purpose (7) and car availability (3). Input: matrices by segment. Input: costs by mode.	Matrices will be very sparse for the captive and competition sub-segments. At this stage it is theoretically feasible to aggregate these over purposes. Need to consider structure of mode choice model (to deal with slow modes and car passengers).

[18] Level-of-service will be at zonal level, so aggregation issues are unlikely to be significant.

10.3 Task 8.3 Observed assignments

There are 2 objectives of this task:

- to obtain an insight into bias and under-reporting in the survey data[19];
- to generate realistic road times for use in model calibration.

Inputs

Car, [CV??] and public transport trip matrices.
Public and private networks.
Screenline counts.
Road speed surveys.
1991-based model estimates for 2001.

Processing

Assign matrices and compare with counts.
Adjust road assignment as necessary. Options:

- scaling of matrix to match counts[20];
- Bayesian combination[21] of survey matrix and 2001 synthetic matrix from 1991-based model;
- use of 2001 survey matrix at some aggregated sector level with zonal disaggregations based on 1991-based model (providing some 'smoothing' of the matrix).

Outputs

A road network with realistic journey times for input to model estimation.
An appreciation of matrix biases, which may effect later workstreams.

Resources

19 Note that we are not collecting an interview screenline which would enable purpose-specific biases to be examined.

20 It is not thought to be necessary to use MVESTM for this; and we ourselves only hold a version for tuning the road traffic matrix.

21 A weighted average based on estimated variances.

10.4 Task 8.4 Calibration

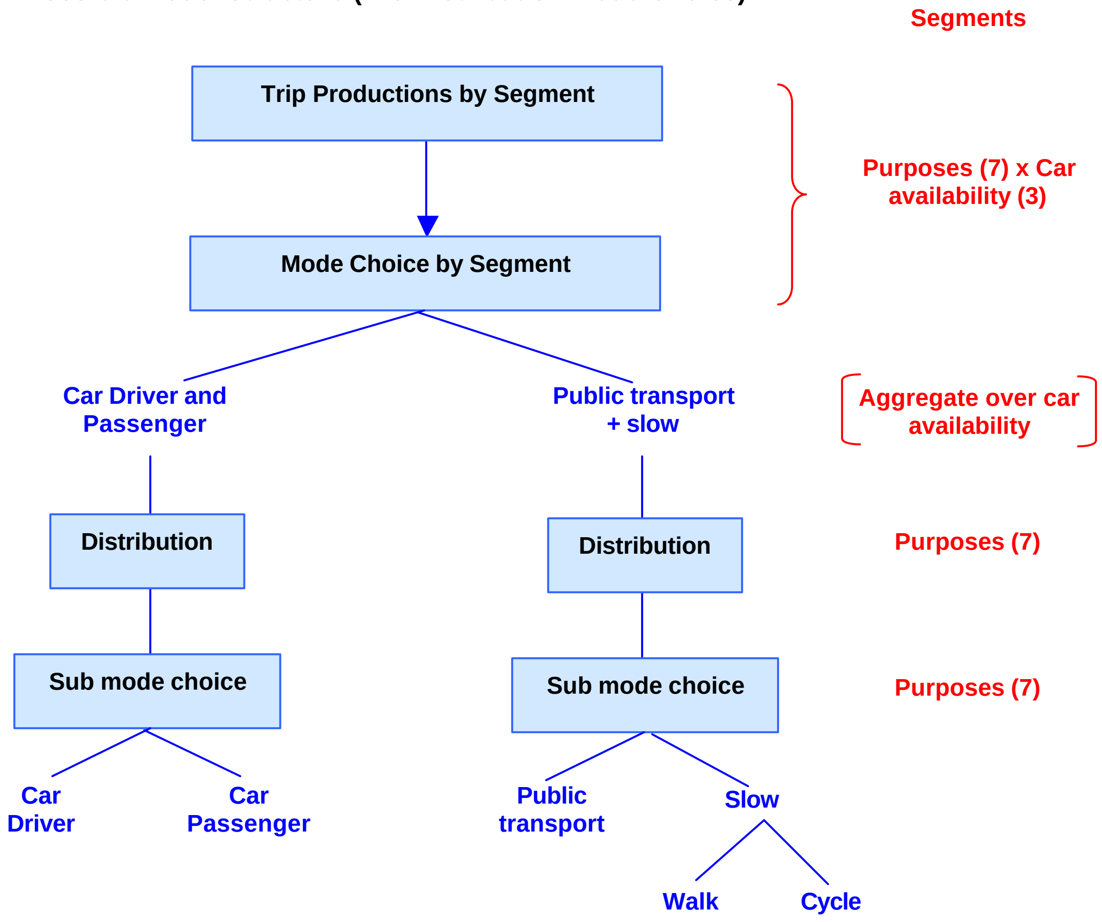
Inputs
Calibration files.

Processing
A programme based on LIMDEP or the preferred software to be prepared.

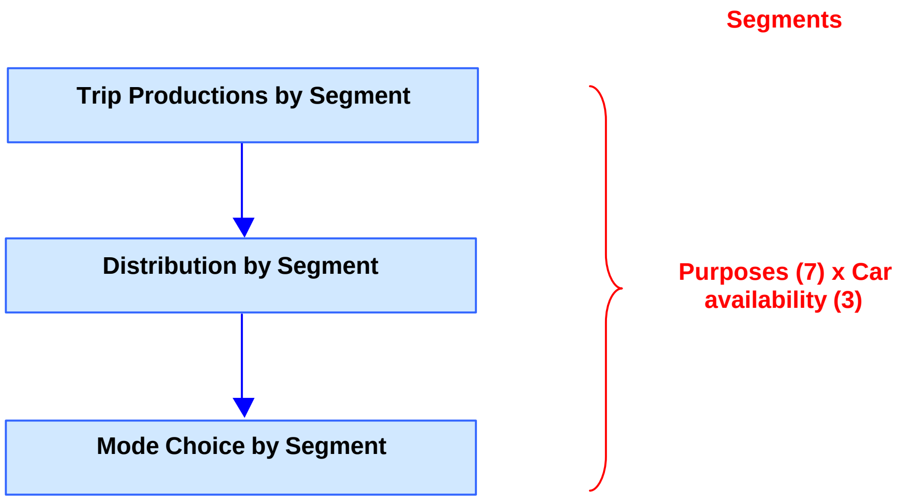
Outputs
Final model.
Report.

Resources

■ **Figure 10-1 Possible Model Structure (Pre-Distribution Mode Choice)**



■ **Figure 10-2 Possible Model Structure (Post Distribution Mode Choice)**



11. Task 9 Commercial Vehicle Forecasting

To be written as a result Task 2.6.

11.1 Task 9.1 Current Year Matrix

11.2 Task 9.2 Growth Factors

12. Task 10 Time Period Factors and Peak-Spreading

12.1 Issues

Principal issues are:

- can we improve on present approach by locating more evidence of peak spreading effects and/or by using latest international research;
- whether a formal link between peak-spreading and the time period factors can be specified.

12.2 Task 10.1 Time Period Factors

Inputs

Observed P/A matrices by purpose, time period and mode.

Processing

The development of factors simply involved computing the proportion of the total matrix for each mode and purpose in each time period, making due allowance for P/A and O/D structures which determine travel direction (as in figure below).

TIME PERIOD FACTORS

24 Hours

P/A

	Res	CBD
Res	20	100
CBD	0	10

=130 trips

From Home OD

	Res	CBD
Res	10	50
CBD	0	5

=65 trips

To Home OD

	Res	CBD
Res	10	0
CBD	50	5

=65 trips

O/D

	Res	CBD
Res	20	50
CBD	50	10

=from home + to home =130

AM Peak

From Home OD

	Res	CBD
Res	8	40
CBD	0	4

=80% from home = 52

To Home OD

	Res	CBD
Res	1	0
CBD	5	0.5

=10% to home = 6.5

O/D

	Res	CBD
Res	9	40
CBD	5	4.5

=from home + to home =58.5

Interpeak

From Home OD

	Res	CBD
Res	1	5
CBD	0	0.5

=10% from home = 6.5

To Home OD

	Res	CBD
Res	1	0
CBD	5	0.5

=10% to home = 6.5

O/D

	Res	CBD
Res	2	5
CBD	5	1

=from home + to home = 13

PM Peak

From Home OD

	Res	CBD
Res	1	5
CBD	0	0.5

=10% from home = 6.5

To Home OD

	Res	CBD
Res	8	0
CBD	40	4

=80% to home = 52

O/D

	Res	CBD
Res	9	5
CBD	40	4.5

=from home + to home =58.5

A further refinement would be to vary the factors by location, but the sample statistics will not permit much of this. It may be worthwhile testing some such segmentations (to, for example, distinguish longer distance work trips to the CBD).

Outputs

Resources

12.3 Task 10.2 Peak-Spreading

An important issue for WRC, the model form used in Wellington is a development of that used in Auckland, itself designed to reproduce a peak-spreading elasticity inferred from the opening of the Sydney Harbour Tunnel. The theoretical structure is not described, so the following develops the model structure.

We assume a logit choice model between the am peak period and the interpeak which uses as measures of utility the generalised cost of travel by car in the two time periods. We use an incremental implementation of the model, pivoting around the base year shares of travel between the two time periods which can be expressed as follows:

$$P^x_{am} = P^0_{am}/[P^0_{am} + (1\text{-}P^0_{am}).\ \exp(\Delta U^x\text{-}\Delta U^0)]$$

where:

P^x_{am}, P^0_{am} are the proportions of am peak trips in the base (0), derived from the time period factors, and alternative (x) scenarios

ΔU^x, ΔU^0 are the differences in the generalised costs of travel GC in the am and midday periods for the two scenarios factored by a sensitivity coefficient which converts them to utilities; ie

$$U^0 = \beta.(GC^0_{mid} - GC^0_{am})$$

where:

β is set to achieve the required elasticity.

In the present model, this module is applied to the total trip matrix and not, for example, by purpose. It is not self-evident that the parameter value is consistent with the model hierarchy, in that it is an order of magnitude smaller than the mode choice and distribution parameters, whereas it should, according to the hierarchical structure, be **larger**.

The effective calibration of a useful model will be difficult. Nonetheless, the interest in investigating road pricing etc, requires some kind of peak-spreading module. Our inclination is towards a simple approach which would provide the required functionality, and in which the coefficients could be guessed/judgemental if necessary.

A key issue for the design is to implement the time-period factors in a way which will be compatible with a subsequent peak-spreading development. This means abandoning the idea of choosing a single period, according to purpose, for the distribution model costs. It might be acceptable to **start** the DMS estimation on such assumptions, but then to switch, ideally to composite costs (over time of day) or, more easily, to average (trip-weighted) costs.

Note that whether we use average costs or time period specific costs as now, the impact of peak pricing with a **fixed** set of time period factors could have the counter-intuitive effect of reducing off-peak travel. This points up the fundamental need for a peak-spreading model to deal with such policies. The present peak-spreading formulation could suffice (though it should be segmented by purpose) provided it can be made compatible with the model hierarchy: if this is not done there is again the possibility of inconsistent model forecasts.

We propose to carry out a brief review of international experience in this area.

13. Task 11 Assignment – Public and Private

Inputs
- Procedures in current model
- VOT by market segment

Processing
Private (road)
Basic approach would be to use a multi-class, generalised cost EMME/2 (equilibrium) assignment. This is a flexible approach which lends itself to improved modelling detail, such as separate assignment by vehicle type and/or purpose (e.g. CVs), and also for different market 'segments' (e.g. can reflect different levels of VOT or 'willingness-to-pay in road pricing/toll studies).

The basic assignment methodology is fixed within EMME/2, but key design decis ions will include:

- Basic user-class segmentation (and possible future segmentations)[22]
- Representation of 'fixed costs' in EMME/2 gc assignment
- Global routeing parameters
- Link-specific routeing parameters (i.e. assignment 'k factors' to account for environmental factors not included in basic gc function, such as strategic signage, comfort, road hierarchy etc)
- Use of separate assignment or pre-loads/fixed flows
- Representation of road-pricing or tolls (most probably through additional fixed-costs on specific links, which also allows differential tolling by market segment when using mode-specific links)
- Sub-segmentation of main modes (i.e. definition of auxiliary modes)
- Network 'priming' ('warm starting')
- Feed-back and update loops, especially in respect of any junction modelling
- Ease of extraction of data required in other modules (e.g. skimmed gc component matrices for use in distribution./mode split
- Convergence
- Running times
- Interface with PT modes (especially ability to extract link speed data for assessing bus speeds)
- Ease of secondary analysis (select-link, sub-area traversals, emission-analysis (cold-start assessments), use of emme/2 Additional Demand Options capabilities (important)

Key input requirements are for routeing parameters. Sources include: PEM, other calibrated models, current Transfund parameter-value research. Even so, there will be a network calibration task to tune the routeing parameter to best fit observed counts.

Public
Key design decisions include:

[22] Different paths for cars and commercial vehicles are probably essential, especially if tolls are considered. For specific tolling studies a further segmentation by income group or by classes of car user affected might be envisaged.

- Representation of generalised cost parameters (fares, penalties, transfers, weights, boarding, ivt, mode-specific weights, 0.75) in the model, and ease of data entry for model runs
- Park & ride – to be considered in Task 2.15
- Slow modes will not be assigned.

Outputs

Resources

14. Task 12 Software Implementation and Interfaces

We need a formal software design process via the preparation of Functional and Technical Designs prior to setting the system up in EMME/2. As far as possible, for efficiency, this should exploit the existing set up. Its should also specify the standards to be applied.

The first, early task will be to prepare an overall system design, so far as possible based on the present set-up, in order that individual modules can be added to it as they are calibrated.

There also some technical elements which will need to be addressed in this, notably:

- how the best estimate matrices developed in Task 1.9 are interfaced with the model system (via some sort of incremental procedures)[23];
- minimising run times will be one of the objectives of the software design;
- the issue of whether the client wishes for the design of specialised user interfaces; this is desirable, but it is not clear how far this can be managed with EMME/2 and how much effort would be involved; this will be reviewed at the outset of the task and, if it looks feasible, a pal will be prepared for discussion with WRC;
- design of convenient standard output facilities;
- a GIS structure to the outputs;
- file naming and calling conventions.

[23] There is experience of this - not all of it good – for the London model, on which we shall need to draw.

15. Task 13 Model Testing and Validation

15.1 Task 13.1 Design of Iterative Procedures

A process of model iteration should be specified and evaluated which leads to convergence of all aspects of the model (essentially the road assignment and demand matrices). The process will involve optimising damping and feedback loops and requires the identification of key convergence diagnostics (eg traffic flows, matrix totals and travel times).

15.2 Task 13.2 Systematic Analysis of Error Transmission

The trip end distribution and mode choice models are individually estimated and fitted to survey information. These are then combined with applications matrices. Individual errors in each model calibration, which may be acceptable at that point, are transmitted to later models and these may have unacceptable consequences. As the models are implemented these errors need to be traced through and if necessary corrected in the earlier models.

15.3 Task 13.3 Specify Test Plan

The process of model testing and the acceptable range of outcomes should be pre-specified. This will include fits to independent data such as counts and the results of model sensitivity tests, such as elasticities.

15.4 Task 13.4 Carry out Testing and Model Validation

In which the test plan is executed.

15.5 Task 13.5 Model Bias Correction and Tuning

The first model tests are likely to show unacceptable errors and this will leads to a period spent searching for errors in the model systems and data – including network errors, errors in observed counts etc etc. In extreme cases it may require recalibration of an individual model. This becomes a substantial issue for the more complex model specifications with more feedback mechanisms, and could lead to the requirement to recalibrate, for example, the accessibility terms in the car ownership model and the distribution, mode choice and peak-spreading models.

16. Task 14 Matrix Estimation

Inputs
Screenline counts.
Model assigned flows on road and public transport.

Processing
The validation may throw up significant model misfits. Some of these may arise from survey biases which may be corrected by uniform factoring, while others may be due to random errors in the model.

Model fit may be substantially improved by using matrix estimation techniques. We have available analytical techniques, based on MVESTM, for the road assignments.

Outputs
Set of matrix adjustment factors.

Resources

17. Task 15 Prepare Base Forecasts

Important Note
There are deficiencies in carrying out this task at present and therefore the detailed procedures for the task will be developed at the outset tom act as a template for later updates.

Inputs
Forecast year networks, planning and economic inputs.
Information on past trends in travel patterns in Wellington.

Processing
Run model.
Check forecasts for reasonableness and validate against past trends (broadly).

Outputs
Forecasts.
Note.

Resources

18. Task 16 Final Reporting

This task involves:
- combining the technical notes into reference material on the model development
- preparation of a User Guide
- preparation of a software implementation reference.

Appendix 1 Task 1 Survey Data and Processing and Other Travel Data

1.1 Common Codes

Inputs
Compatibility equivalence table.
Survey data base (note this is applied to all surveys, not just the household survey).

Processing
Specify variables affected.
Set up common codes in data base (eg the aggregated model purpose codes; aggregated model mode codes, captive choice definitions etc)[24].
Attach zone codes, station codes??

Outputs
Updated survey data base.
Report.

Resources

[24] Specifically concerned with defining NHB classes.

1.2 Household Survey Trip Linking

Inputs
Household survey data base.

Processing
Specify the trip linking process. In drawing up the specification, it is wise to look carefully at the data to ensure that the trip-linking specification is reasonably comprehensive. The specification should also include the extraction of diagnostic data to check the trip-linking and to capture oddities. This is likely to involve:

- Identifying sequences of trip stages that form a single trip.
- Converting the trip stage records into a trip record. Typically, this will include resetting:
  - the end time to that of the last stage,
  - the destination purpose to that of the last stage,
  - the mode of travel, where more than one mode is used for the stages, based on the combination of a modal hierarchy and the mode used for the greater part of the trip (see below),
  - parking cost, from the appropriate final car stage,
  - for car trips, the number of passengers, using some appropriate logic where the numbers vary between the stages,
  - etc.

  Specify the main mode of transport, as follows:
  - specify a modal hierarchy (eg ferry, rail, bus, car driver, car passenger, cycle, walk);
  - also calculate the crow-fly distance travelled by each transport mode;
  - compute main mode based on both the hierarchy and the distance travelled;
  - cross-tabulate and identify proportion of data where the main mode is different for the 2 tests;
  - check a sample of trips and develop ‘rules’ where these tests are inconsistent;
  - specify final mechanism for identifying main mode.

Implement trip-linking.

Outputs
A trip file in the survey data base.
Note.

Resources

1.3 Household Survey MIC

Inputs
Survey data base.
MIC report.

Processing
Identified MIC issues are:
- missing travel diaries
- missing key items (there should be no missing values):
  - household: size, number of cars
  - person: age, employment status, work address, study address
  - stage: start and end addresses, start and end times, purpose, mode
- other missing items.

In Task 1.3, the frequency of occurrence of all MIC issues will have been established. For missing travel diaries, three options might be considered:
- factoring up the other household trips
- developing factors by person characteristics, which would be applied as overall factors for the whole sample
- applying a correction factor to the interview clusters (of 7 households).

In principle, a combination of the last 2 methods is preferred based on the following correction factors:
- compute the number of persons out of each cluster 'c' in age categories 'a' (eg 6-21, 21-64 full time employed, 21-64 other, >64): N_{ca}
- compute the number of missing diaries are missing in each cluster by age: n_{ca}
- compute correction factor to be applied to all trip records: $1/(1- n_{ca}/ N_{ca})$

This factor ensures that the locational characteristics of the missing travel are not lost but also is sensitive to the idea that some persons are more likely to be missing than others.

Tasks are therefore as follows.
- Review occurrence of missing diaries.
- Specify bias correction factor.
- Implement and verify results.

For missing values for key items, it is only necessary to confirm that none are in fact missing.
For "other items", review occurrence of missing values.
Specify action to take on each (either code as missing, or substitute artificial data, the process of estimation to be specified).
Implement.

Outputs
Trip Correction Factors in data base.
Synthesised values for some questions in data base.
Note.

Resources

1.4 Household Survey Expansion

Inputs
Household survey data base.
2001 census households by zone.
2001 census household structure (number of persons/no of full time employed persons) by "combined TLA"[25].

Processing
We must cover expansion and bias correction.

Specify process, along following lines.

1. Decide on the geographical areas to form the basis of expansion:
   - in principle there should be a minimum sample in each area (say 30 household interviews); with 2500 interviews this suggests about 100 areas;
   - also check sample frame (design of survey) and ensure that stratifications/sampling structures are systematically reflected in expansion process.
2. Count the sample of surveyed households in each area, compare with the census and compute the ratio, the sample expansion factor; review range of values before proceeding.
3. Table the expanded sample data (by sub-region, say local government area) by household type (family size * number of full time employed persons[26]); aggregate table where sample is small (<30 households - a judgement is needed here); compare with planning data distributions and compute bias correction factors; review range of values before proceeding.
4. Re-do task (2) this time applying bias correction factors to sample prior to re-computing sample expansion factors.
5. Then overall expansion factor for each element of the sample is the product of the expansion factor and the bias factor.

Implement.

Outputs
Expansion factor for each household in the data base.
Report.

Resources

[25] In which Masterton, Carterton and S. Wairarapa are combined.
[26] Commonly it is found that response rates vary by household size and number of employed persons, simply because of difficulty of intercepting respondent(s).

1.5 External 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
MCCs
ATCs

Processing
For interviews this involves:

- defining MIC (only questions 5 & 10 can be excluded);
- range checks on questionnaire responses;
- rejecting questionnaires failing MIC;
- if possible, checking on reasonableness of address data.

For counts it involves:

- MCC: checking for completeness and miscoding;
- ATC: this involves checking for tube failure and missing data.

Outputs
Cleaned roadside questionnaire data base
Cleaned MCC & ATC files

Resources

1.5.2 Processing

Inputs
Cleaned roadside questionnaire data base
Cleaned MCC & ATC files

Processing
Survey period expansion factors are to 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 may vary by vehicle type, given low proportion of CVs);
- aggregate MCC counts by the 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.

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, it is suggested that survey day bias is corrected by also applying a factor based on the survey day 24 hour traffic volume to the average weekday traffic volume in the survey week(s).

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

Out of scope data needs to be excluded which, in this case means identifying and eliminating residents' travel (which is already encompassed by the household survey:

- the preferred method is to process the external data to eliminate residents' home-based trips, identify non-residents' home based trips, and isolate non home based trips; the latter could be either by residents or non-residents and will need to be factored down to exclude residents' trips; an approximate set of factor(s) would be the residents/non-residents split for the home based external trips by time of day;
- the matrix of residents trips should be retained in case we think it useful to bring it back into the model to improve the synthesised matrix.

The data base then needs to be set up to be exactly compatible with the household survey. *These tasks need to be defined but may include:*

- recoding the purpose codes in the shorter roadside questionnaire to the codes to be used in the modelling, and to be consistent with the household survey;

The roadside interviews were in both directions of travel so no trip reversal is required.

1.6 Train and Bus Survey Editing and Processing

To be verified on final survey specification

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

1.6.1 Cleaning and editing interviews and counts:

Inputs

Rail and bus questionnaire data base
Interviewer hand-out records
ETMs

Processing

For interviews this involves:

- defining MIC (questions 4, 6, 9, 10a & b, 17-19 can be excluded from MIC; if this leads to high rejection rates, we may reconsider the status of 11-16);
- range checks on questionnaire responses;
- rejecting questionnaires failing MIC;
- if possible, checking on reasonableness of address data.

For interviewer records it involves: checking for consistency, completeness and miscoding.

We assume that the ETM information provided is accurate.

Outputs

Cleaned questionnaire data base
Cleaned interviewer records

Resources

1.6.2 Processing

Inputs

Rail and bus survey questionnaire data base.
Interviewer records
ETM period data

Processing

Survey period expansion factors are to be appended to the trip records as follows:

- pre-process questionnaire data base, using the questionnaire number in conjunction with interviewer records, to allocate a station of survey to each record and a time of survey (according to the time periods recorded by the interviewer);
- tabulate interview samples at each station by time period; determine whether to use 15 min, 30 min, 1 hour or longer basis for the expansion period (this decision will vary by station and by peak/off-peak based on usage); in this process new issues may arise such as stations with no achieved questionnaires in a time period, which may require special adjustment;
- aggregate interviewer counts by station 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 station and time period;
- append expansion factors to interview records;
- verify process by tabulating expanded data and comparing with counts and/or assigning to network.

Do we need to think about bias in self-completion surveys (mainly the bus/rail surveys)?

We know of no double-counting nor any out-of-scope data.

Trip reversal. The survey of inbound trips (actually, 'boarding a southbound train') needs to be reversed to obtain the outbound matrix. For complete data flexibility, I propose below to duplicate the trip records, but this is not absolutely necessary and simpler process could be devised:

- duplicate records of every respondent; then amend the duplicated records as follows:
- delete duplicates where response to Q11 is No;
- swap the answers to Q1 and Q7, Q2 and Q8, Q5 and Q9, Q4 and Q6;
- process Q12-Q14 to give the new trip start trip time, to go in Q3; note start times may be outside our direct survey period (either before or after)
- retain Q10a & b, Q15-19;
- *we shall need to consider some additional correction factors – we need to correct for the deleted duplicates otherwise we shall end up with fewer trips outbound than inbound – this is essentially a uniform factor applied to all retained duplicate questionnaires; we also need to consider the implications of the duplicated trips outside the survey period .*

Is 24 hour expansion needed?

The data base then needs to be set up to be exactly compatible with the household survey. *These tasks need to be defined but may include:*

- recoding the purpose codes to the codes to be used in the modelling, and to be consistent with the household survey;

Outputs

Resources

1.7 School Survey Processing/Pre-Analysis

Inputs
School survey data base (cleaned).
List of schools giving role, and identifying years covered.

Processing
Because of the novelty of the survey and the structure of the data, expansion will not initially be attempted.

The processing will first be concerned with understanding how behaviour varies by school year (ie pupil age) in regard to mode of transport, trip length and time of travel. It will be sensible to segment these analyses by broad geographical area (say local authorities).

This analysis will be an input to consideration of the education model in Task 2.7.

The data will allow the direct development of an observed trip matrix, but it is likely that customised techniques will be needed to exploit the detail in the data. These will need to be developed with an understanding of the data and the approach to modelling education trips, but may be based on such concepts as:

- use the school rolls as zonal school population controls;
- allocate pupil origins, modes and times of travel based on likelihoods developed from the surveys;
- control these to planning data on population by age.

This type of technique will allow the infilling of missing data. It may be appropriate then to combine the synthesised matrix with the observed matrix using a Bayesian weighting technique.

Outputs

Resources

1.8 Annualisation

Inputs

Given our main concern is car traffic and public transport usage, we probably should seek:

- continuous traffic count data,
- ETM data for bus services,
- equivalent data for rail services.

The exact nature of the data requirement will depend on the exact timing of the surveys.

Processing

At this stage we merely note the general requirements/principles. All of the surveys occur in some limited period of the year, from a 2 month period for the household survey to specific weeks for the other surveys. While the periods are intended to be 'neutral', it would be best to verify this, for which we need some form of data collected throughout the year. This data can be used to expand the survey samples to represent annual averages.

Outputs

Resources

Appendix 2 Task 7 Sample Enumeration

For example, what is done in START is along the following lines:

For any zone, find the vector of household type proportions $\mathbf{h} = (h_1, h_2, \ldots. h_{13})$ such that:

(1) $$\sum_{t=1}^{13} h_t = 1$$

(2) $$\sum_{t=1}^{13} h_t . a_t^p = \pi^p \ , p = 1{,}4$$, where π^p = average no. of persons type p per household [p = children, unemployed adults, employed adults, retired]
and

(3) $$h_t \geq 0, \forall\, t$$

The values of the matrix $\mathbf{A} = \{a^p{}_t\}$ give the (average) number of persons of type p in household type t: these may be assumed fixed. The πs are the input "targets" based on the MERA population forecasts. may change markedly. It is assumed that the total number of households in the zone is also known.

Ignoring the non-negativity constraints, this is a problem in 5 equations and 13 variables, so it is under-identified. The solution proposed is to minimize some function of deviations from the base household proportion template $\mathbf{h}^*$ subject to the constraints in Eqq 1 – 3. An obvious choice is the sum of squares:

$$\min z(\mathbf{h}) = (\mathbf{h} - \mathbf{h}^*)^T(\mathbf{h} - \mathbf{h}^*) = \sum_{t=1}^{13}(h_t - h_t^*)^2$$

Without the non-negativity constraints, this can be solved by standard Lagrangean multiplier techniques: with a little luck, the optimum might satisfy condition (3) without further effort. Otherwise, we are in the realms of non-linear programming. We can still work with the Lagrangean, but we have to deal with possible boundary solutions.

Since the minimand is quadratic and strictly positive, it is a convex function, while the five constraints are all linear in the problem variables $\mathbf{h}$. This is a standard problem, and has a unique solution. Writing the Lagrangean multiplier for condition (1) as λ and the multipliers for condition (2) as ϕ^p, the first order conditions can be written:

(a) $$h_t . \left(\frac{\partial z}{\partial h_t} - \lambda - \sum_p \phi^p a_t^p \right) = 0 , \forall \text{ t}$$

(b) $$\frac{\partial z}{\partial h_t} - \lambda - \sum_p \phi^p a_t^p \geq 0 , \forall \text{ t}$$

(c) $$\sum_t h_t = 1 \text{ and } \sum_t a_t^p h_t = \pi^p , \forall \text{ p}$$

(d) $$h_t \geq 0 , \forall \text{ t}$$

This is a quadratic programming problem, and code can be written to solve this, without significant difficulty. There are other variants of the process which relax the equality constraints, while attaching different levels of "importance" to meeting them.

Appendix 3 Model Simplifications

Introduction

The present model technical specification is quite complex in regards to the family structure model and the car ownership model. It could be argued that the resources involved are disproportionate compared to their value in the model when car ownership levels are high in Wellington and growth is expected to be small[27] and when family size is not expected to change significantly.

The model specification requires a forecast of person trips classified by some measure of captivity to public transport. Person trip volume for non-work trips is expected to be somewhat greater for persons in more mobile (ie car owning households), a difference found in the existing model's trip rates, while captivity is intended to be related to the availability of a car.

The model is proposed to be structured almost entirely on the basis of person trips. It is the need to represent car ownership in the model which necessitates a household dimension to the model and this in turn requires the development of a family structure model able to relate the characteristics of persons and households.

What lies behind this proposal is the thought that, if we could find a means of representing car ownership effects at a person level, much simplification might follow.

The Trip End Model

Its is expected that the models for non-work trips will have a mobility effect on the trip rates. As presently specified, the model form would be:

HBSh/HBSo trip rate = $T_i + \Delta Tj$
where:
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.

Thus there is:

- a trip rate for each person type
- plus an incremental trip rate, uniform across the person types, dependent on the car ownership of the household.

[27] Booz Allen forecast a 13% growth nationally from 2001 to 2021).
The current model has the following household car ownership forecasts:

	1996	2016
0 cars	14%	10%
1 car	50%	51%
2+ cars	36%	39%

For example, for adults the trip rate might be 0.5 for 0 car households, 0.6 for 1 car households, 0.65 for 2+ car households etc, where the car ownership-related trip increments of 0.1 and 0.05 respectively would be the same for all person types.

Suppose we re-specified this model as:

HBSh/HBSo trip rate = $T_i + \alpha.CA$
where:
T_i is the average trip rate for person type i,
CA is the cars/adult in the household (a convenient measure of mobility, as we shall see)
and α is the 'mobility' coefficient.

This model is a simplification of the original model in that it assumes that the relationship with the number of cars is linear (ie with twice as many cars/adult the mobility trip increment is doubled, an assumption not made in the original specification). But given that we expect this effect to be second order, such an approximation seems quite reasonable.

Captivity/Choice Trips

In the original specification, persons of each type were to be grouped into households of different car ownership levels, and their trips for each trip purpose accumulated and split between captive, competition and choice categories.

Suppose we define captive, competition and choice on the basis of the ratio of cars to adults in the household in the following way:

	Cars/adult
captive:	0
competition:	0>cars/adult<1
choice:	1

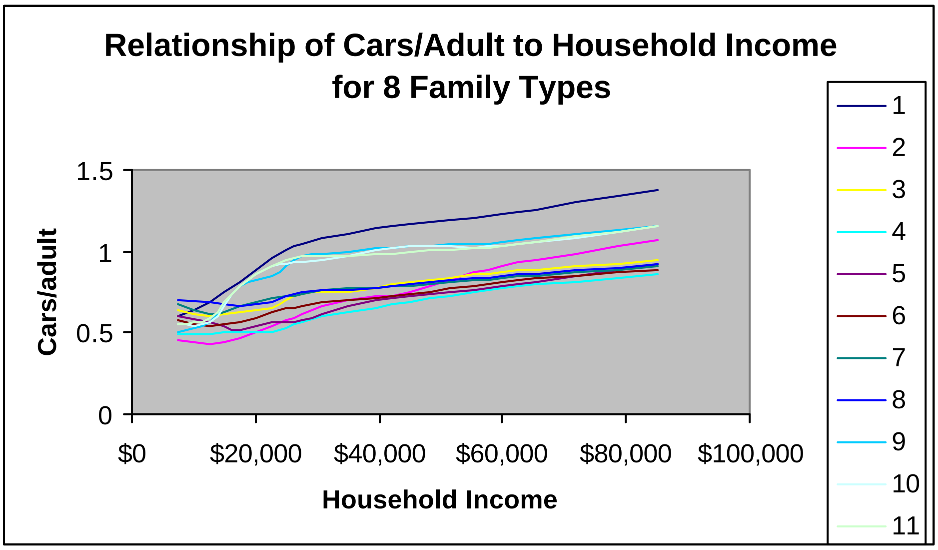
Forecasting Car Ownership

Work in the UK and elsewhere has suggested that household car ownership model forecasts should be constrained to some independent overall trend control. Booz Allen & Hamilton carried our work for Transit in 1997, which provides a basis for such forecasts[28]. The forecasts are of cars/person, but these are easily adjusted to give cars/adult using the MERA current and forecast distributions of the population by age group.

The figure below is an analysis of trends in cars/adult with income for 11 family types[29].

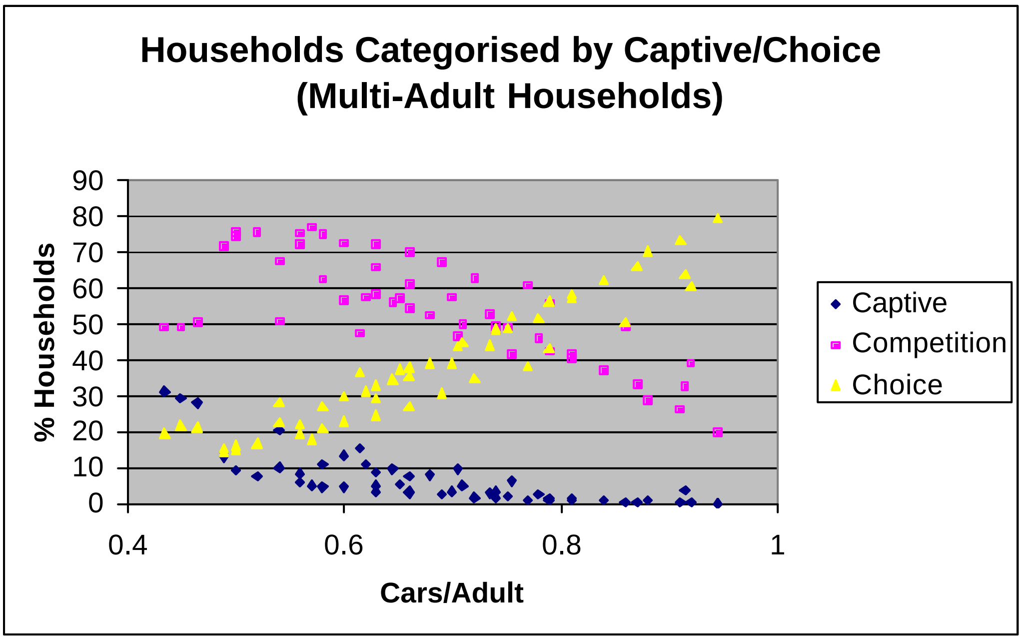
The figure tends to suggest that as income increases the cars/adult for each family type tends to increase also, and by not dissimilar amounts. Indeed, except for a few outlying family types the relationships are very similar[30]. That is, using this measure of mobility, the approximation of applying a uniform trend in cars/adult to all household types does not seem unreasonable.

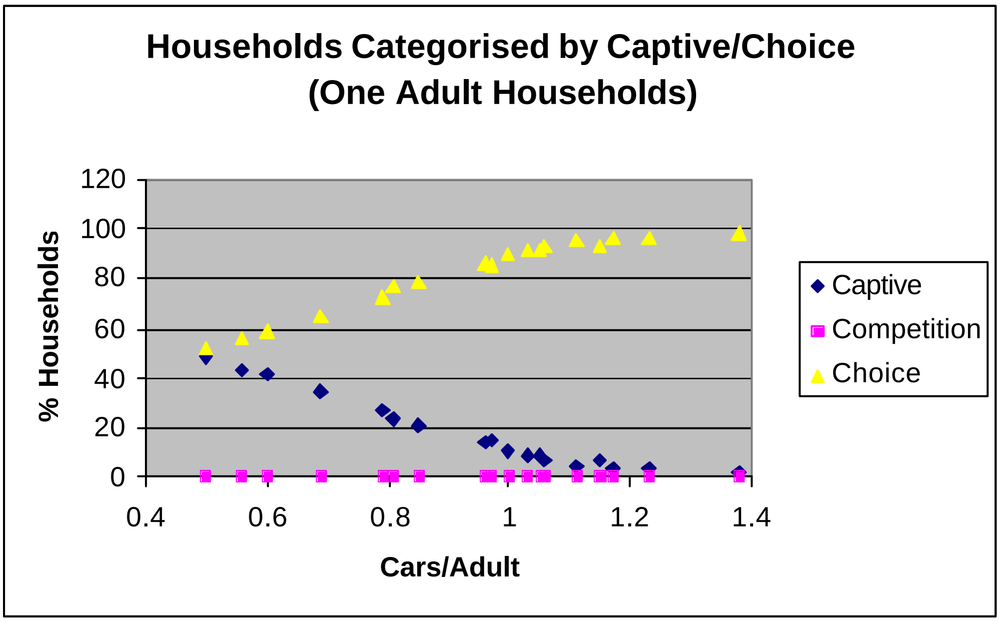
[28] Vehicle Availability Forecasting Model, August 1997.
[29] The analysis is approximate, the precise number of adults in some family types being unknown.
[30] Outliers: high cars/adult: single people and one parent families with dependent children.



Relationships Between Cars/Adult and Captivity

Again using the BAH data, the figures below show how the proportions of households in captive, competition and choice categories varies as the average cars/adults increases. Separate relationships apply for 1 adult households. It is clear that reasonably consistent relationships apply, with the number of choice households increasing and the other categories decreasing as cars/adult increases.





Putting Together these Concepts into a Revised Model Structure

The concept is reasonably straightforward.

The BAH model, or some further development of it, will provide the basis of the car ownership model. It will forecasts cars/person and, by adjustment, cars/adult – call this CA^y, y being the forecast year; a value for the base year will be available from the census. A household-based model will not be developed.

Trip end models will be developed on a person basis as planned. Household attributes will not be included except the mobility-related attribute 'household cars/adult' for the social, shopping and recreational trip purposes. That is:

person trip rate = $T_i + \alpha.CA$

where:

T_i is the average trip rate for person type i,

CA is the cars/adult in the household (a convenient measure of mobility, as we shall see)

and α is the 'mobility' coefficient, non-zero only for shop, recreational and social trips.

MERA will provide forecasts of the number of persons by zone N^y_{iz} in the categories i in future year y. So estimates of zonal trip ends for each trip purpose (subscript not shown) will be:

$T^y_z = \Sigma_i N^y_{iz}(T_i + \alpha.CA^y_z)$, where CA^y_z will be estimated as described below.

Households in the interview sample will be classified into cars/adult categories. For these aggregated categories, and for each purpose separately, person trips will be aggregated into the captive, competition and choice categories, to give relationships of the form illustrated immediately above. It is probable that single person households will need to be distinguished and also that the resulting curves will need to be smoothed for implementation. So for each purpose, we shall be able to relate the proportion of trips in each of the categories captive, choice and competition to the household average cars per adult.

Using census data, every zone can be set a base year mean cars/adult value – CA^{2001}_{z}. In forecasting this value will be increased uniformly, according with the regional, overall car ownership forecast to give CA^{y}_{z} for each future year y.

Given the zonal cars/adult, the aforementioned calibrated curves enables, for each purpose, the split of the zonal trip ends into captive, choice and competition segments for each zone. In doing this some projection or scenario estimate of the proportion of 1 person households will be needed.

The above process requires no household car ownership model and no complex family structure model.

Comment

I identify below the key approximations made in this process as a check of the risks:

- we assume essentially uniform car ownership growth across the zones in the region – *the justification for this approximation is (1) the expected low car ownership growth and (2) the generally poor performance of car ownership models at reproducing zonal growth variations anyway*
- we do not allow household structure variables in the person trip end models except for a ‘mobility’ terms which is a linear function of cars/adult – *all of this seems broadly reasonable as we will be picking up the main person type effects*
- we use an aggregated method of allocating the captive, competition, choice split for the trip ends based on a relationship with cars/adult calibrated from the household survey; we then apply this at a zonal level – *this is the fundamental new relationship and is a risk if it turns out that the relationships which I have illustrated above are not so clear cut in the data.*

I sought a second opinion on the concept from John Bates in the UK. His response is as follows.

“The only nervousness I have is that associated with departing from conventional procedures - not because I'm basically conservative, but because every time you change something in this game there may be some consequence down the line that you've never seen before. So I think caution is valuable. Also, though I can't see anything wrong with the simplifications you're proposing (and the attached spreadsheet of UK trip rates (all modes) generally bears you out - NB this has been hastily put together based on other people's analysis, so it's not assumption-free data, but I think it's generally OK), I hadn't seen the co/hh structure model as being as resource-consuming as you. On balance, I would have stayed with it, but taken a sharp tool to any attempts at introducing new developments.

But I assume that you are worried about something in the cost estimates, and want to cut it down. Apart from the above nervousness, I can see no reason why it shouldn't work just as well as the conventional approach (which, as you say, isn't that brilliant anyway). None of the risks you identify look serious: the greater risk of the unknown, and having not thought something out early enough."

Recommendation

Unless there are strong in principle objections, we conclude that this simplified approach deserves serious consideration, having the potential to reduce costs and simplify the model fairly substantially.

Alternative Simplification

If it is felt necessary to retain a car ownership model, the following approach to family structure would appear to offer considerable prospects for reducing the cost of developing this sub-model.

The method is illustrated in the tables over the page. A base year cross-classification of person types by household types is obtained from the household survey data (the "base distribution").

For the forecasts years, zonal forecasts by person type (MERA) and household car ownership forecasts (from the model) are obtained and using the formula shown used to synthesise the cross-distribution (the "synthesised forecast distribution). This method has wide applicability.

Base Distribution

	Persons									
	Infant	Child	young adult unemployed	employed	adult unemployed	employed	adult >60 unemployed	employed	Households	*(Total Persons)*
Households by Type										
Car ownership = 0									0.15	*0.1*
Retired, 1 adult	0.0025	0.0075	0.0050	0.0025	0.0050	0.0150	0.0100	0.0025		*0.05*
Non-retired, 1 adult	0.0013	0.0038	0.0025	0.0013	0.0025	0.0075	0.0050	0.0013		*0.025*
Retired, 2 adults	0.0006	0.0019	0.0013	0.0006	0.0013	0.0038	0.0025	0.0006		*0.0125*
Non-retired, 2 adults	0.0006	0.0019	0.0013	0.0006	0.0013	0.0038	0.0025	0.0006		*0.0125*
3+ adults	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0*
Car ownership = 1									0.35	*0.3*
Retired, 1 adult	0.0050	0.0150	0.0100	0.0050	0.0100	0.0300	0.0200	0.0050		*0.1*
Non-retired, 1 adult	0.0050	0.0150	0.0100	0.0050	0.0100	0.0300	0.0200	0.0050		*0.1*
Retired, 2 adults	0.0025	0.0075	0.0050	0.0025	0.0050	0.0150	0.0100	0.0025		*0.05*
Non-retired, 2 adults	0.0025	0.0075	0.0050	0.0025	0.0050	0.0150	0.0100	0.0025		*0.05*
3+ adults	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0*
Car ownership = 2									0.35	*0.35*
Retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0*
Non-retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0*
Retired, 2 adults	0.0050	0.0150	0.0100	0.0050	0.0100	0.0300	0.0200	0.0050		*0.1*
Non-retired, 2 adults	0.0050	0.0150	0.0100	0.0050	0.0100	0.0300	0.0200	0.0050		*0.1*
3+ adults	0.0075	0.0225	0.0150	0.0075	0.0150	0.0450	0.0300	0.0075		*0.15*
Car ownership = 3									0.15	*0.25*
Retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0*
Non-retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0*
Retired, 2 adults	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0*
Non-retired, 2 adults	0.0050	0.0150	0.0100	0.0050	0.0100	0.0300	0.0200	0.0050		*0.1*
3+ adults	0.0075	0.0225	0.0150	0.0075	0.0150	0.0450	0.0300	0.0075		*0.15*
(Total	*0.05*	*0.15*	*0.1*	*0.05*	*0.1*	*0.3*	*0.2*	*0.05)*		

Synthesised Forecast Distribution

	Persons									
	Infant	Child	young adult unemployed	employed	adult unemployed	employed	adult >60 unemployed	employed	**Input Household Forecasts**	*(Check on Total Persons)*
Households										
Car ownership = 0									**0.05**	
Retired, 1 adult	0.0004	0.0015	0.0012	0.0003	0.0015	0.0049	0.0045	0.0008		*0.01505*
Non-retired, 1 adult	0.0002	0.0008	0.0006	0.0002	0.0008	0.0024	0.0023	0.0004		*0.00753*
Retired, 2 adults	0.0001	0.0004	0.0003	0.0001	0.0004	0.0012	0.0011	0.0002		*0.00376*
Non-retired, 2 adults	0.0001	0.0004	0.0003	0.0001	0.0004	0.0012	0.0011	0.0002		*0.00376*
3+ adults	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0.00000*
Car ownership = 1									**0.3**	
Retired, 1 adult	0.0019	0.0077	0.0062	0.0015	0.0077	0.0252	0.0232	0.0039		*0.07742*
Non-retired, 1 adult	0.0019	0.0077	0.0062	0.0015	0.0077	0.0252	0.0232	0.0039		*0.07742*
Retired, 2 adults	0.0010	0.0039	0.0031	0.0008	0.0039	0.0126	0.0116	0.0019		*0.03871*
Non-retired, 2 adults	0.0010	0.0039	0.0031	0.0008	0.0039	0.0126	0.0116	0.0019		*0.03871*
3+ adults	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0.00000*
Car ownership = 2									**0.4**	
Retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0.00000*
Non-retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0.00000*
Retired, 2 adults	0.0026	0.0103	0.0083	0.0021	0.0103	0.0335	0.0310	0.0052		*0.10323*
Non-retired, 2 adults	0.0026	0.0103	0.0083	0.0021	0.0103	0.0335	0.0310	0.0052		*0.10323*
3+ adults	0.0039	0.0155	0.0124	0.0031	0.0155	0.0503	0.0465	0.0077		*0.15484*
Car ownership = 3									**0.25**	
Retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0.00000*
Non-retired, 1 adult	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0.00000*
Retired, 2 adults	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		*0.00000*
Non-retired, 2 adults	0.0038	0.0151	0.0120	0.0030	0.0151	0.0489	0.0452	0.0075		*0.15054*
3+ adults	0.0056	0.0226	0.0181	0.0045	0.0226	0.0734	0.0677	0.0113		*0.22581*
Input Person Forecasts	**0.025**	**0.1**	**0.08**	**0.02**	**0.1**	**0.325**	**0.3**	**0.05**		

Formula

$$P'_{tc} = P_{tc} *(H'_c/H_c) * P'_t / \Sigma_c(P_{tc}*H'_c/H_c)$$

P: person, t: type
H: household; c: car ownership level
‘ refers to future year