B3 - Modelling Modal Choices with Discrete Choice Models: An Application to the North Shore Busway

Yi-Min Huang and Judith Wang[1]
yi_min_h@hotmail.com

Abstract

The Northern Busway Project in Auckland will be completed in early 2008. The project involves five new bus stations each incorporates park and ride facilities, and a dual-lane busway being built alongside the Northern Motorway. This busway aims to improve the bus service to improve the modal share of buses and in turn improve the traffic on the Northern Motorway. The project costs approximately 300 millions dollars (Transit NZ, 2007) and therefore it is important that its impact on the future travel demand be forecast as accurately as possible. The forecasting model used for applying for funding for the Northern Busway is based on generalised cost, which takes into consideration only monetary cost and time, to forecast the travel demand. This type of model did not recognise many factors which cannot be mathematically measured, such as comfort and personal preferences, but crucial for the mode choice decisions made by travellers.

The purpose of this study is not to reforecast the usage of the North Shore Busway but to illustrate how a different approach can be applied to produce more accurate forecasts for similar kind of projects in the future. In this research, discrete choice model is used which allows the travel demand model to incorporate these immeasurable factors on a disaggregate basis. Data collection procedure and instruments were designed to obtain both Revealed Preference (RP) and Stated Preference (SP) Data. The data was then used to estimate the RP and SP models using Maximum Likelihood Method which is performed by LIMDEP in this research.

In conclusion, it appeared that trip chaining activity is a significant factor encouraging travellers to use private vehicles, while qualitative reasons such as convenience and comfort are the main factors encouraging travellers to use public transport. This is because their existence causes travellers' perception of the importance of other factors, such as cost and time, to reduce. It is important that SP experiments are designed to simulate the actual decision process as much as possible to avoid bias in the results.

[1] Lecturer, Department of Civil and Environmental Engineering, The University of Auckland

Introduction

The Northern Busway Project in Auckland consists of five bus stations and the busway structure with three motorway interchanges to improve bus accessibility for the North Shore. The stations are, from north to south, Albany, Constellation, Sunnynook, Smales Farm (Formally Westlakes) and Akoranga. The busway structure is being constructed along the southbound of Northern Motorway where each station is linked either directly to the motorway (Albany station) or to the busway. At the time of this study, Albany and Constellation stations had been completed and were operational while buses were travelling using the shoulder of the Northern Motorway as the busway was still being constructed. Once the busway structure is completed, it will initially be used only by buses, emergency vehicles and service vehicles. It is also proposed to allow limited number of High Occupancy Vehicles (HOVs) to use the busway from Constellation station.

The modal split in the demand model for the Northern Busway Project developed in 2000 for economic evaluation was estimated by an incremental Logit model based on generalised cost function (Andjic *et al.*, 2001). The generalised cost function follows the conventional approach which includes only monetary cost and travel time. Modal choice decisions, however, as we all aware are much more complex which can also be affected by comfort, convenience, personal preferences and etc.

This paper aims to demonstrate the process of developing a Discrete Choice Model with which more aspects of the modal choice decisions can be considered. It is a disaggregate model that is based on factors considered by individual travellers. It originated from behaviour psychology and has been used increasingly throughout the world for the last 30 years (McFadden, 2001). There are two case studies in New Zealand in the literature, one by Laird and Nicholson (1994) studying mode choice of students travelling to the University of Canterbury and another study on the commuter behaviour in New Zealand's three largest cities, namely, Auckland, Wellington and Christchurch (Hensher *et al.*, 2004).

Literature Review

This section describes the background of Discrete Choice Models and the methodologies as related to data collection and data analysis.

There are two types of data collection methods for Discrete Choice Models: Revealed Preference (RP) and Stated Preference (SP). RP method is based on the existing travel behaviour while the SP method is based on responses of hypothetical scenarios designed by the modeller. Each method has its pros and cons. As the RP method collects information on the existing decisions, it also captures the behaviour and perception of the travellers. On the other hand, the data are likely to contain measurement errors and has a strong correlation between factors (e.g. cost and time) (Kroes and Sheldon, 1988). Flexibility is the main advantage of the SP method, as the experiment is designed entirely by the modeller. The main weakness of SP method is the reliability of the survey results, that is to say whether the stated behaviour represents the actual behaviour. This discrepancy has been identified to result from fatigue of respondents, policy response and alternative bias (Bates, 1988). The first can result from asking each respondent to answer too many scenarios. Policy response bias occurs when respondents intentionally bias the answer in an attempt to influence the model results. When a respondent has un-intentionally over-valued the existing choice, the alternative bias occurs. These effects can be minimised or prevented in the design of the SP experiment.

Discrete choice models in general evaluate the probability of individuals to choose a particular alternative which is "*... a function of their socioeconomic characteristics and the relative attractiveness of the options*" (Ortuzar and Willumsen, 2002). The general form of the discrete choice model can be described by the following formula:

$$P_n(i) = \frac{U_n(i)}{\sum_C U_n(c)}. \qquad \text{(E.q. 1)}$$

where

$$U_n(i) = V_n(i) + \varepsilon \text{, and} \qquad \text{(E.q. 2)}$$

$$V_n(i) = \sum \beta f(x_i, S_n). \qquad \text{(E.q. 3)}$$

In words equation (1) means that the probability of individual *n* choosing alternative *i* equals the proportion of the utility (*U*) (and hence the relative attractiveness) of alternative *i* over the sum of utility of all the alternatives in the choice set *C* which includes alternative *i*.

Equation (2) shows that the utility function consists of two parts: a systematic component (*V*) which is determined by the modelling process and an error component (ε). The error term is assumed to have zero mean and varies according a certain type of probability distribution function. Equation (3) further defines the systematic component which equals the sum of decision attributes in relation to the alternatives (x_i) and socioeconomic attribute in relation to the individual (S_n). β is the coefficients associated with each attributes which reflects the relative importance of the attributes on the systematic utility.

The probability distribution function which ε takes on will determine the type of discrete choice model. In practice, the most commonly used discrete choice model is Multi-Nominal Model (MNL) which assumes ε is Identically and Independently Distributed (IID) Gumbel Type I Extreme Value. This means that MNL assumes that all alternatives are independent of each other and all individuals have the same taste, i.e. the probability of choosing alternative *n* is independent and identically distributed among individuals, which results in the following formula:

$$P_n(i) = \frac{e^{\mu V_{ni}}}{\sum_{C_n \in C} e^{\mu V}} \qquad \text{(E.q. 4)}$$

where μ is the scaling factor of the Gumbel Distribution while $C_n \in C$ denotes that the individual choice set (C_n) is a subset of universal choice set (C). It is commonly used in practice because it can be estimated conveniently without using any approximation methods. It can be estimated using Maximum Likelihood Method. There are other models which relaxes the independent alternatives and identical tastes assumptions, since these can sometimes cause counter-intuitive results if the choice context does not fit the assumptions. For example, Nested Logit Model groups the alternatives that are dependent of each other together, while groups are independent of each other. Heteroscedastic Extreme Value Model, on the other hand, is a model that allows for difference in taste among individuals.

Maximum Likelihood (ML) Method is used to estimate the β values as shown in Equation (3), It is based on the notion that a sample is more likely to be drawn from a particular population, even though it could have been drawn from several different populations. Its general form is:

$$L(\theta) = \prod_{n=1}^{N} \prod_{i \in C_n} P_n(i)^{g_n(i)} \qquad \text{(E.q. 5)}$$

where $L(\theta)$ is the ML function, $P_n(i)$ is the probability of individual *n* choosing alternative *i* and $g_n(i)$ equals 1 if the individual *n* choose alternative *i* while equals 0 otherwise. In this study, this estimation process was performed by LIMDEP which is a software package with readily available statistical estimation techniques to suit most experimental needs.

Methodology

Before defining the universal choice set (C), it is necessary to make assumptions on other aspects of travel choice as follows:

(i) Travellers will always choose to travel as communal trips are non-voluntary.
(ii) Travellers' destinations will be any destination south of the Harbour Bridge.
(iii) Travellers' arrival time is before 10 am.
(iv) Travellers will commute using the Northern Motorway rather than other routes.

As the main focus of the study is related to mode choice, the universal mode choice set contains five alternatives. These are Low Occupancy Vehicle (LOV), High Occupancy Vehicle (HOV), Park and Ride (ParkRide), Kiss and Ride (KissRide) and Walk and Ride (WalkRide). The former two are private vehicle alternatives where low occupancy is defined as less than three travellers in a vehicle while high occupancy means otherwise. The latter three are public transport alternatives with the difference in the method of access to the bus station. Motorcycles have been excluded as it is expected that the influence of the busway on this mode will be minimal as motorcycles can travel through the motorway between the traffic. Moreover, Cycle and Ride is excluded because it has minimal users.

Individual Choice Set (C_n) is defined here as the set of mode choices that is available to the individual n from the Universal Choice set (C). It is therefore dependent on the characteristics of the individuals (e.g. car ownership, driver's licence, etc) and this information is to be captured in the survey.

The next step is to define the socioeconomic (S_n) and decision (x_i) attributes in Equation (3) and these are tabulated in Table 1.

<table>
<tr><td>Alternatives</td><td colspan="7">Decision</td><td colspan="3">Socioeconomic</td></tr>
<tr><td></td><td>Petrol Cost</td><td>Parking Cost</td><td>Fare Cost</td><td>Access Time</td><td>Waiting Time</td><td>Parking Time</td><td>In-Vehicle Time</td><td colspan="3"></td></tr>
<tr><td>LOV</td><td colspan="2">LOV Cost</td><td rowspan="2"></td><td>LOVacc</td><td rowspan="2"></td><td rowspan="2">CarPark</td><td>LOVveh</td><td rowspan="2">TripChain</td><td rowspan="2" colspan="2"></td></tr>
<tr><td>HOV</td><td colspan="2">HOV Cost</td><td>HOVacc</td><td>HOVveh</td></tr>
<tr><td>ParkRide</td><td>PB Cost</td><td rowspan="3"></td><td rowspan="3">Bfare</td><td rowspan="2">VBacc</td><td rowspan="3">Bwait</td><td rowspan="3"></td><td rowspan="3">Bveh</td><td rowspan="3"></td><td rowspan="3">Income</td><td rowspan="3">Quasum</td></tr>
<tr><td>KissRide</td><td>KB Cost</td></tr>
<tr><td>WalkRide</td><td></td><td>WBacc</td></tr>
</table>

Table 1. Decision and socioeconomic attributes.

The decision attributes are listed within Table 1 under the Decision column with rows showing their related alternatives and columns correspond to information the attribute contains. For example LOV Cost attribute is related to LOV alternative and contains petrol and parking cost. Some attributes are shared among some alternatives because it is assumed that the attributes are the same across the corresponding alternatives, such as bus fares are the same regardless how the patrons arrive at the bus station. The socioeconomic attributes listed here are the only three that were included in the final models as others are either insignificant or cannot be estimated. Other socioeconomic attributes considered were gender, age and availability of vehicle, driver's licence, company vehicle and parking. TripChain attribute refers to the need for some travellers to use their private vehicle during work or after work which acts as an incentive to drive. Quasum is the total number of qualitative reasons that a respondent gave for using the public transport.

The questionnaire for the RP survey contains three pages: The first page is about their background information and socioeconomic characteristics, the second page asks for information on private vehicle modes (LOV and HOV) while the last page is for public transport modes (ParkRide, KissRide and WalkRide). Questions on age and income were optional. Diagrams were drawn to aid the understanding of the questions in pages two and three.

Factor	**High Level (1)**	**Low Level (0)**
LOVcost	Increase by 30%	Decrease by 30%
HOVcost	Increase by 30%	Decrease by 30%
Bfare	Increase by 30%	Decrease by 30%
KBcost	Increase by 30%	Decrease by 30%
PBcost	Increase by 30%	Decrease by 30%
LOVacc	Increase by 5 minutes	Decrease by 5 minutes
HOVacc	Increase by 5 minutes	Decrease by 5 minutes
WBacc	Increase by 5 minutes	Decrease by 5 minutes
VBacc	Increase by 5 minutes	Decrease by 5 minutes
Bwait	Increase by 5 minutes	Decrease by 5 minutes
CarPark	Increase by 5 minutes	Decrease by 5 minutes
LOVveh	Increase by 30%	Decrease by 30%
HOVveh	Increase by 20%	Decrease by 20%
Bveh	Increase by 20%	Decrease by 20%

Table 2. Attribute Levels for each Attribute

For the SP survey, the experiment was designed using factorial design. The attribute levels for each attribute are shown in Table 2 as the high (1) and low (0) levels. Each level is derived as a relative change from the RP data instead of a fix value to make the scenarios more realistic. Since each respondent will be provided with scenarios that were based on his/her existing experience. The scenarios, at the same time as being realistic, will also need to change significant enough from the existing experience to cause possible change of choice. 30% was recommended by Hensher *et al.* (2005) as a reasonable compromise between the two criteria. For access, waiting and parking time, 5 minutes were used because 30% changes of these values would be insignificant. To ensure that HOV and bus in-vehicle time do not exceed LOV in-vehicle time, it was decided to use 20% changes instead. As there are a total of 14 decision attributes, a fractional factorial design with 16 factors were adopted as in Sloane (2006) with 32 treatments. This design is for the estimation of main-effect only and hence it is assumed that there is no interaction between the attributes. The two extra factors were used as blocking factor to divide the treatments into 4 blocks of 8 treatments. Furthermore, two unrealistic treatments were taken out to prevent bias in the results.

A notebook computer was used to generate the scenarios during the interview in the SP survey. The respondents' RP responses were first verified. Since in the RP survey, each respondent was asked to provide information on two modes which corresponding to the existing choice and the next-best choice. When a respondent's first and second choices are both private vehicle mode or both public transport mode, a third choice/scenario would be requested which had to be associated with the other mode. Then the SP scenarios were shown using Excel one by one. Each respondent was asked "If you are to undertake the same trip as this morning, given the alternatives listed here with the associated attributes, which one will you choose to use?" to frame the choice context for all the scenarios.

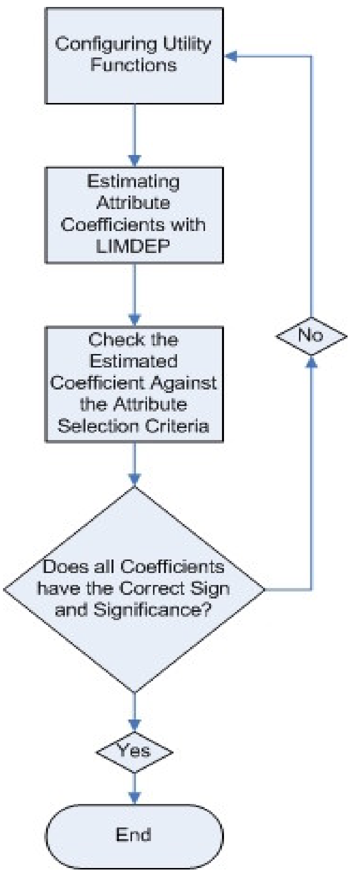

Figure 1. Analysis process

The data analysis process is illustrated in Figure 1. The coefficients in the systematic utility function were estimated using LIMDEP. The Attribute Selection Criteria at the third step is listed in Table 3.

Parameters		Attributes	
		Choice	Socio-Economic
Correct Sign	Significant	Include	Include
	Not Significant	Include	May Reject
Wrong Sign	Significant	Big Problem	Reject
	Not Significant	Problem	Reject

Table 3. Attribute Selection Critera

The aim of the analysis process is to find the best systematic utility function by altering the configuration of the attributes (x_i, S_n) and the coefficients (β), as defined in eqn (3), for each set of data. Decision attributes were mostly fixed and cannot be altered except for access and parking for private vehicle where in one of the utility function these two was combined. Socioeconomic attributes, on the other hand, are included in the utility function only if its effect appeared to be statistically significant. For each attributes in the utility function there must be an associated coefficient, but a coefficient can be shared among a number of attribute across the alternative if necessary. For example, VBacc and WBacc may share the same coefficient if they are deemed to have the same effect on travellers' behaviour.

Data Summary and Results

For RP data, 75 data points were obtained out of 106 respondents. 10 data points were from staff and students in the University of Auckland, 20 from the users at the Albany station and 45 were collected on the Northern Express buses. Other survey methods were tried, such as interviewing at shopping centres and random distribution of flyers to households, but failed to collect usable/representative samples due to low response rate. As a result, the sample collected is biased heavily towards public transport as only 15% of the respondents choose to use private vehicle. The age span between 21 and 25 is the largest sampled group constituting 23% of the sample. The highest income category was NZ$ 70,000+ which constitutes 25% of the sample. These respondents appeared to be more time sensitive than the others - one of these respondent explicitly said that the ability to perform work during the journey is very valuable. Most of the respondents' household have at least one vehicle per driver. A total of eight different qualitative reasons were collected for reason to use public transport; these are comfort, convenience, reliability, environmental friendliness, service quality, dislike of traffic on the motorway (risk/stress aversion), the ability to perform work while travelling, and the ability to socialise with friends/co-travellers.

140 data points were collected in the SP surveys from 20 respondents. The characteristics of the SP data are similar to that of the RP data as mentioned above. One point stands out in both the RP and SP data is that none of the respondents has opined that HOV is feasible for them. Some of the respondents share a private vehicle with one other passenger but that is still categorised as LOV. Although the sample collected is not representative of the North Shore population, it indicates that allowing HOV onto the busway might have limited effect on the traffic condition on the motorway overall.

Ideally, a utility function for each mode choice should be estimated. Unfortunately, due to the lack of data, only two sets of utility functions (the RP and SP models) were estimated, with two alternative modes: private vehicle (LOV data) and public transport (with data from all three bus alternatives). Details of the estimated utility functions are illustrated in Table 4.

Attributes	**RP model**				**SP model**			
	Coefficient	Standard Error	t-ratio	P[|Z|>z]	Coefficient	Standard Error	t-ratio	P[|Z|>z]
Private Vehicle								
Monetary Cost	-0.311	0.164	-1.895	0.058	-0.221	0.052	-2.602	0.009
In-vehicle Time	-0.130	0.081	-1.601	0.110	-0.036	0.022	-0.973	0.331
Total Access Time	-0.067	0.114	-0.586	0.558				
Access Time					-0.125	0.056	-1.371	0.170
Parking Time					-0.09	0.071	-0.815	0.415
TripChain					4.186	0.973	2.644	0.008
Public Transport								
Constant (ASC)	1.255	3.153	0.398	0.691	-0.229	1.342	-0.105	0.917
Monetary Cost	-0.183	0.404	-0.454	0.650	-0.063	0.124	-0.313	0.754
In-vehicle Time	-0.100	0.085	-1.175	0.240	-0.061	0.022	-1.206	0.228
Access Time	-0.223	0.118	-1.895	0.058	-0.043	0.039	-0.676	0.499
Waiting Time	-0.184	0.255	-0.721	0.471	-0.085	0.058	-0.899	0.369
Income	-0.049	0.027	-1.835	0.067				
Quasum	3.198	1.301	2.458	0.014				
Log Likelihood Ratio	-9.166				-41.377			

Table 4. Final RP and SP systematic utility functions.

The coefficient refers to the β value in Equation (3). The t-ratio is a statistical value to assess whether a coefficient is statistically significant. P[|Z|>z] represents the probability of wrongly rejecting the null hypothesis that the coefficient is zero. The log likelihood ratio at the bottom of the table represents the quality of the model, specifically, the larger the value the better the model. However, as the RP and SP data came from essentially different samples in this research, the comparison of this value between the two models is not meaningful. The ASCs (Alternative Specific Constant) for public transport models are dummy attributes which represent the relative attractiveness of the alternatives comparing to a reference alternative, which in this case is the private vehicle. Specifically, the ASC captures all the residual effect that is not captured by the decision and socioeconomic variables. Therefore in an ideal situation with all effects taken into account, ASC should equals to zero.

In private vehicle models, access and parking time have been combined as total access time for the RP model while the SP model retains the two time attributes separately. The reason is that the estimated coefficient of parking time for the RP model had a positive sign, which is a counter-intuitive answer as increase in time spent shall result in reduction in utility. The coefficient of ChainTrip attribute for the RP private vehicle model and the coefficient of Income and Quasum of the SP public transport model could not be estimated due to lack of data.

The public transport ASC for the RP model has a positive sign while in the SP model it has a negative sign. This means that in the RP model public transport is preferred over private vehicle and vice versa for the SP model. There are a number of explanations for this; one is that there are other socioeconomic attributes relating to public transport in the RP model that had been omitted or that the Quasum attribute is not a good representative qualitative attribute for public transport. Looking at the cost coefficients, it can be seen that travellers put more weight on cost when considering using private vehicles; as shown in both the RP and SP models, as the cost coefficients of private vehicle are both larger than that of public transport. Comparing the time coefficients, the total access time coefficient is smaller than the in-vehicle time coefficient. This is unusual as it is expected that the latter coefficient will be much greater than the former, as shown in the RP model where the access time coefficient is 2.3 (0.223/0.100) times the other. This may be caused by measurement errors as access time of private vehicle are usually unperceived by travellers. In the SP private vehicle model, by showing the access time explicitly in the interview, the sensitivity to access time has greatly increased (access time is approximately 3.5 times (0.125/0.036) in-vehicle time).

To compare the effect of socioeconomic attributes with time and cost, the utility function will need to first be converted into the same unit, say, utils. In the analysis, the coefficients were used to compute the average contribution (*C*) of each attribute in the total magnitude of change of the utility function. Algebraically, let u_{jk}' be the total magnitude of change of utility of data points *j* when Quasum or TripChain is at *k* level while β_i and x_{ijk} are coefficients corresponding to attribute *i* and attribute values corresponding to attribute *i* of data point *j* that has *k* level for Quasum or TripChain attribute respectively. (Note in particular that, each data point corresponds to a set of attribute values as related to one alternative provided by one respondent.) Then

$$u_{jk}' = \sum_i \left|\beta_i \times x_{ijk}\right|. \qquad \text{(Eq. 6)}$$

The contributions of magnitude of change of attribute *i* in the utility of each data point *j* that as *k* level for Quasum or TripChain attribute is written as

$$C_{ijk} = \frac{\left|\beta_i \times x_{ijk}\right|}{u_{jk}'}. \qquad \text{(Eq. 7)}$$

Finally, the average contribution of the magnitude of change for attribute *i* for all data points *j = 1, …, J* that has Quasum or TripChain attribute level of *k* equals

$$\overline{C}_{ik} = \frac{\sum_{j=1}^{J} C_{ijk}}{J}. \qquad \text{(Eq. 8)}$$

$\overline{C}_{ik}$ was calculated for all attributes at all the Quasum and TripChain attribute levels.

These are expressed in percentage in Table 5. The first column of the table lists all the levels that each of Quasun and TripChain have taken in the data set. The second column shows the number data points corresponding to each Quasum or TripChain level. In addition, $\overline{C}_{ik}$ which is shown as the percentages in the above table will be referred to as simply the contribution to the utility, or just contribution.

From the top part of the table, it can be seen that as Quasum level increases (from 0 to 4), the contribution to the utility of Quasum increases from 0 to 53% while the contribution of other attributes decreases. When Quasum is at level 1, the contribution of the Quasum attribute is almost as significant as in-vehicle time and access time. The contribution of the Quasum attribute then becomes the largest across all the attributes when it takes on the level of 2. This shows that travellers become less sensitive to other attributes such as time and cost when considering their travel choices as they have more qualitative reasons to use the public transport.

For the TripChain attribute in the SP private vehicle model, a similar conclusion can be drawn. Considering the bottom part of Table 5, as the TripChain level increases (from 0 to 1), its contribution inclines by 38% while the contributions of other attributes (In-vehicle and access time) decline sharply except for parking time and monetary cost which remain almost the same. From this result, it can be concluded that the need to undertake trip chaining activities has dominant impact on travellers' travel choice as it reduces the relative importance of in-vehicle time and access time greatly.

RP Public Transport Model – Quasum Attribute								
Quasum Levels (*k*)	Data Points	ASC	Cost	In-vehicle Time	Access Time	Waiting Time	Income	Quasum
		(%)	(%)	(%)	(%)	(%)	(%)	(%)
0	32	10	8	24	31	9	18	0
1	35	8	8	19	23	7	14	21
2	23	6	6	15	21	4	14	33
3	1	6	6	16	11	3	15	44
4	1	5	5	13	9	4	11	53
SP Private Vehicle Model – TripChain Attribute								
TripChain Levels (*k*)	Data Points		Cost	In-vehicle Time	Access Time	Parking Time		TripChain
			(%)	(%)	(%)	(%)		(%)
0	252		40	24	33	4		0
1	18		38	15	5	4		38

Table 5 Contribution of Magnitude of Change Utility by each Attribute for RP public transport and SP private vehicle model.

Discussions

The main advantage of Discrete Choice Models is its flexibility. This is demonstrated in this research by the Access Time for the RP private vehicle model and the use of the Quasum variable as an attribute. Furthermore, if any interaction between the attribute is of interest, then it can be done by altering the factorial design. The type of discrete choice model to be used can also be changed. MNL is a convenient but restrictive model while other more sophisticated models would allow the decision behaviour to be modelled in more detail at the price of more information needed. Another advantage is that its ability to model non-existence alternatives, although this was not demonstrated in this research due to data shortages.

On the other hand, the main shortcoming of discrete choice model is its need for massive information that requires person-to-person interviews. This specifically refers to the SP data collection method. For RP data collection, it is possible to ask the respondents to simply fill out questionnaires without direct interviews with good questionnaire design. To collect the information needed using SP method, large amount of resources will need to be committed. From hiring the interviewers, to training the interviewers, to advertising request for interviewees, to performing the interviews with its associated needs for space and time, to quality control of the interviewing process and lastly to record all the information.

A number of techniques for the RP and SP surveys can be identified from this research. First of all, if the RP questionnaire is to be answered without an interview, a figure depicting the meaning of each question will decrease the amount of unusable responses. Also, it is important to minimise the number of pages of the questionnaire. For the SP survey, it is important that the interviewing instruments represent the real decision making process as much as possible; as shown in this research, explicitly showing the access time for private vehicle affects the resulting utility function.

Conclusions and Suggestions for Future Research

From the analysis of the final models, a number of conclusions can be drawn about the behaviour of the travellers and the effects of the data collection instruments. Firstly, Travellers with qualitative reasons, such as convenience and less stress, are less sensitive to other attribute such as cost and time. Secondly, the need to undertake trip chaining activities during or after work has a similar effect to the qualitative reasons to the private vehicle utility. The need to undertake trip chaining activities, especially, greatly reduce travellers' perception on the importance of access time. On the other hand, it has little effect on the perception of parking time and monetary cost.

As discussed above, the main advantage of discrete choice model is its flexibility from experimental design to choice of model. While the main disadvantage is its need for large amount of resource for careful data collection, in particular if SP survey is conducted.

To extend this research, future research can be conducted to investigate the individual effect of the qualitative reasons as opposed to the total effect of a number of qualitative attributes. This shall improve the understanding of travellers' behaviour and help the council and/or bus companies to improve the ridership of the buses. Research efforts can also be put on devising a method to distinguish KissBus and ParkBus as they essentially share the identical cost and time characteristics. Another area open for future research is investigating the effect of the SP interview instrument on the travel behaviour. Finally, with enough resources, the research methodology in this study can be applied to develop a full model to forecast the travel demand on North Shore Busway. Such model can be incorporated into the four-stage strategic planning model for the Auckland Region to improve the accuracy in demand estimation.

References

Andjic, Z., Foy, B. A., *et al.* (2001). The North Shore Busway Evaluation.

Bates, J. J. (1988). "Econometric issues in stated preference analysis." Journal of Transport Economics and Policy **XXII**(1): 59 - 69.

Hensher, D. A., Greece, W. H., *et al.* (2005). Applied Choice Analysis: A primer. New York, Cambridge University Press.

Hensher, D. A., O'Fallon, C., *et al.* (2004). "Constraints affecting mode choices by morning car commuters." Transport Policy **11**(1): 17-29.

Kroes, E. P. and Sheldon, R. J. (1988). "Stated Preference Methods: an introduction." Journal of Transport Economics and Policy **XXII**(1): 11 - 26.

Laird, J. J. and Nicholson, A. J. (1994). The University of Canterbury Travel Survey 1993: A Stated Preference Approach to Modelling Modal Split. Christchurch, The University of Canterbury.

McFadden, D. (2001). Disaggregate Behavioural Travel Demand's RUM Side: A 30 years retrospective. Travel Behaviour Research: The Leading Edge. D. A. Hensher. Oxford, Elsevier Science Ltd.

Ortuzar, J. d. D. and Willumsen, L. G. (2002). Modelling Transport. Chichester, John Wiley & Sons Ltd.

Sloane, N. J. A. (2006). "A Library of Orthogonal Arrays." Retrieved 28th October, 2006, from http://www.research.att.com/~njas/oadir/index.html#HAD.

Transit NZ. (2007). "Northen Busway - Costs." Retrieved 11 September, 2007, from http://www.busway.co.nz/index.php/Costs.