GPS User Preferences Exploratory Work

Report for the

New Zealand Ministry of Transport

Prepared by

Dr Richard Meade
Cognitus Advisory Services Limited

with

Dr Lydia Cheung
AUT University

4 July 2016

cognitus

Contents

1 Introduction and Summary

1.1 Purpose of the Research

This research investigates the feasibility of using existing New Zealand travel survey data in discrete choice analysis (DCA) to understand transport users' preferences for various transport attributes. In particular, the New Zealand Ministry of Transport (the Ministry) has 11 years of historical travel data from its Household Travel Survey (HTS). It wishes to see if, and how, that data – perhaps combined with other data – might be used in DCA to estimate how transport users make trade-offs between transport attributes. The attributes in question are:

- Travel time;
- Reliability/variability of travel time;
- Safety;
- Ride quality; and
- Cost of travel.

The ultimate purpose of estimating such trade-offs is to use them as inputs in analyses relating to activity classes in the Ministry's Government Policy Statement (GPS). Those activity classes include:

- State highway investment and maintenance;
- Local roads investment and maintenance;
- Regional improvements and Auckland;
- Public transport;
- Walking and cycling;
- Road policing and safety promotion; and
- Investment management.

1.2 Approach

1.2.1 Discrete Choice Analysis

DCA is an empirical methodology with economic foundations and interpretation.[1] It enables estimation of how decision-makers trade off different attributes of various choice alternatives, when those alternatives are mutually exclusive, exhaustive, and finite.[2]

When one of the attributes in question is measured in monetary units (e.g. travel costs), the trade-off between that attribute and some other non-monetary attribute (e.g. travel time) represents the decision-maker's willingness-to-pay (WTP) for that other attribute. In other words, it represents how the decision-maker trades off units of that other attribute for money.[3] Since transport users' WTP for any non-monetary transport attribute is expressed in monetary terms, this enables direct comparison of transport users' trade-offs between non-monetary attributes.[4]

Even without a monetary attribute variable, DCA enables estimation of how decision-makers trade off non-monetary attributes for other non-monetary attributes.[5] For example, it is possible to estimate how many extra units of travel time users are prepared to incur in order to secure an extra unit of safety, however time and safety are measured.

Since the purpose of estimating transport-users' trade-offs for transport attributes is to inform analyses of GPS output classes, it is preferable to reduce such trade-offs to monetary terms. Hence, estimating WTPs is the focus of this research, for which at least one monetary transport attribute is required.

In this study, we adopt the following naming conventions:

- Value of Time (VOT) – WTP for travel time savings;
- Value of Reliability (VOR) – WTP for more reliable travel (i.e. less dispersion in travel times, however measured);
- Value of Safety (VOS) – WTP for transport safety (whether that safety

[1]For a succinct overview, see Chapter 2 of Small and Verhoef (2007). More extensive introductions are provided in Hensher et al. (2015), and Train (2009).

[2]Train (2009).

[3]Holding the decision-maker's overall satisfaction level, or "utility", constant.

[4]For example, if transport users' WTP for safety is \$3 per unit of safety, but WTP for ride quality is \$2 per unit of quality, then transport users are prepared to trade off $3/2 = 1.5$ units of quality for each additional unit of safety.

[5]Formally, DCA enables estimation of decision-makers' "marginal rate of substitution" between different attributes.

is defined in terms of fatality risk, accident risk, or risk to person or property from crime); and

- Value of Quality (VOQ) – WTP for transport quality.

1.2.2 Suitable Data Types

The methodology can be applied with either or both of two types of data. The first is so-called revealed preference (RP) data, based on decision-makers' actual choices. The HTS falls into this category, as it is a survey of transport users' actual travel choices. The second is "stated preference" (SP) data, usually obtained by presenting survey respondents with hypothetical choice situations where the choice alternatives are carefully structured to enable production of statistically significant estimates of DCA model parameters. Often times RP is combined with SP data to provide richer datasets and enable estimation of more subtle economic relationships between variables of interest.[6]

An advantage of RP data is its realism, since it is based on actual choices. However, it can suffer from other problems, such as inadequate variation in the data to enable precise statistical estimation. Conversely, SP data suffers from the risk that choices, being hypothetical rather than actual, might not reflect how decision-makers will behave in reality. It has the advantage, however, of giving the researcher greater control over the nature of data obtained, and is essential when exploring preferences for choice alternatives which might not be available in RP data (e.g. for novel choice alternatives).[7]

1.2.3 Using DCA to Estimate Transport-Related Trade-Offs

DCA using RP data is therefore an appropriate methodology to consider when exploring how HTS data might be used to estimate transport users' trade-offs between the attributes listed above. More particularly, DCA can be used to estimate how transport users trade-off each of the four non-monetary transport attributes for the one monetary attribute, cost of travel. This then estimates those users' WTP for travel time savings, improvements in reliability (or lowered variability) of travel time, ride quality, and safety.

The approach in this study has been to:

[6] See Table 5 of Li et al. (2011) for a summary of transport reliability studies using either or both data types.

[7] Hensher (2010) discusses hypothetical bias in SP studies, and how it can affect WTP estimates. He provides evidence that SP studies often produce results differing significantly from those in RP studies.

- Determine the current best practice in transport-related DCA studies for estimating transport users' WTP, for the attributes of interest, using RP data;
- Examine what data the HTS contains that can be used to implement suitable DCA methodologies for estimating those WTPs;
- Identify gaps or limitations in the HTS data that might be remedied using other data sources in order to enable DCA to be used to estimate those WTPs; and
- Recommend how the HTS data – combined with additional data – could be used to estimate transport users' preferences for the attributes of interest using DCA.

As part of determining current best practice in transport-related DCA studies, a particular focus has been on identifying:

- What sort of transport choices have been considered – e.g. choice of travel mode, trip purpose, or travel route; and
- How transport attributes have been defined.

Also, in making recommendations as to how HTS combined with other data can be used to estimate transport users' preferences, particular attention has been paid to which type of DCA is appropriate. It is well-documented that different types of travellers make quite different trade-offs between transport attributes. Significantly, a large part of the variation in WTP across users relates to user or choice alternative characteristics not observed by the researcher.[8] Hence it is necessary to consider DCA approaches that produce not just point estimates of WTP for each attribute, but also distributions of estimates – allowing for taste variations across transport users due to both observed and unobserved characteristics of users and transport choice alternatives.

Finally, this research also refers to how estimates of WTPs for the transport attributes of interest can be used in transport analyses relevant to the GPS activity classes referred to above.

1.3 Summary

1.3.1 Application of DCA to Estimate Transport User Trade-Offs

By way of summary, DCA provides a suitable framework for estimating transport users' trade-offs between different transport attributes:

[8] For example, see Small (2012).

- DCA methods are very well established for estimating VOT, having been used for this purpose since the 1970s.[9]
- In the past 10–15 years, methodologies have also become well-developed for estimating VOR separately from VOT.[10]
- More recently, DCA research has emerged on how to estimate VOS separately from VOT, measured in terms of risk of fatal accidents.[11]
- DCA studies also often include quality attributes, as relevant explanatory variables when estimating VOT, VOR or VOS. However, quality attributes are also used in DCA for directly estimating VOQ, though typically for motorised vehicles only.[12]

Typically transport DCA models include VOT and travel cost, one of VOR or VOQ, and attributes of travel alternatives which can include quality dimensions. In principle all such measures can be incorporated in the same DCA model, enabling simultaneous estimation of VOT, VOR, VOS and VOQ.

Since it is well-documented that WTPs for transport attributes vary significantly across different transport users, it is necessary that any DCA allows for observed and/or unobserved differences across transport users. In transport studies, this is typically achieved through the use of a particular version of DCA – the mixed logit (ML) model.[13]

1.3.2 Applicability of Estimated Transport User Trade-Offs in Transport Analyses

In terms of the applicability of estimated user trade-offs for GPS activity classes:

- Estimates of VOT have been used in transport sector analyses since the 1970s.[14]

[9]See, for example, McFadden (1974) regarding the seminal application for the Bay Area Rapid Transit (BART) system in San Francisco. For surveys and meta-analyses, see Small (2012), Hensher (2011), Abrantes and Wardman (2011), and Shires and de Jong (2009).

[10]For surveys, see Carrion and Levinson (2012), and Li et al. (2010).

[11]See Steimetz (2008), Hensher (2006). These studies emphasise the trade-off between lower speeds (i.e. greater travel times) and higher accident rates.

[12]For example, Espino et al. (2006) compare bus and car travel, including quality measures.

[13]Small (2012), Small and Verhoef (2007).

[14]McFadden (1974) discusses the use of VOT in forecasting demand for a novel transport system.

- Increasingly, both VOT and VOR are applied in transport project evaluations and policy analyses, such as the range of evaluations set out in NZTA's Economic Evaluation Manual (NZTA (2016)).[15]

- VOS derived from DCA studies is now also used in transport evaluations, such as for Australian toll-roads.[16]

- VOQ from DCA studies has been used in constructing indices of public transport service quality.[17]

1.3.3 Using HTS and Other Data to estimate Transport User Trade-Offs

The HTS contains much useful data for estimating transport users' trade-offs between the transport attributes of interest using DCA. This includes:

- Overall trip travel times and lengths, with defined start and finish times, days of week and months of year, as well as data required to chain-link trips into journeys;

- Mode choices – car, bus, cycle, walking, etc;

- Some data suitable for measuring ride quality – e.g. vehicle age, size, etc;

- Trip purposes; and

- Traveller demographics – i.e. place of residence, age, personal income, work/study/retirement status, etc.

However, the HTS lacks certain basic data, such as direct travel costs, and details of transport users' non-chosen transport alternatives (i.e. it contains data on only chosen transport alternatives). Thus, in order for the HTS data to be useful for estimating even just VOT, it will be necessary to augment the HTS using other data sources.

There are a variety of data sources already identified which might usefully supplement the HTS in order to enable DCA estimation of VOT etc. These include:

---

[15]Indeed, the EEM includes estimates of both VOT and VOR, including those derived using DCA – Vincent (2008). For a survey of the use of VOR in transport policy and project applications internationally, see OECD/ITF (2010).

[16]Hensher (2006).

[17]For example, see Hensher (2015), Mazulla and Eboli (2006).

- NZTA data on direct travel costs by various modes;
- Public transport databases regarding public transport availability and cost;
- Published accident and crime data useful for measuring transport safety; and
- Public records on certain measures of vehicle quality for public transport, such as vehicle type and fleet age, which can be used to proxy for ride quality.

Supplementing the HTS with data from these other data sources should enable estimation of the required user trade-offs using DCA.

Finally, the HTS contains too few data on long-distance and business-related trips. Hence it is more suitable for applying DCA in relation to short-distance and non-business related trips.

1.4 Recommendations

We recommend that:

- The HTS be supplemented with data from other sources such as those indicated above, to enable DCA to be used to estimate transport users' trade-offs for the transport attributes of interest;
- A DCA model of dominant mode choice for journeys (i.e. chain-linked trips) be estimated, including measures of direct travel cost, travel time, travel time dispersion, safety and quality as explanatory variables to enable estimation of VOT, VOR, VOS and VOQ;[18]
- The model be restricted to urban rather than long-distance travel, and to personal (including work) rather than business-related trips;
- A mixed logit model be estimated, using traveller-level data (including demographics), to estimate how WTPs for transport attributes vary by type of transport user;
- The model be estimated using the full traveller-level dataset, with explanatory variables included for journey location (e.g. urban, rural, etc), and separate models estimated if those explanatory variables indicate significant differences by location; and

[18] Other relevant explanatory variables include, for example, journey length, purpose, timing and location.

- Various measures of travel safety and ride quality be tried, since their measurement is less settled in the literature, and perhaps only for motorised vehicles.

If estimation of a mode choice model using disaggregated data should prove problematic, it is also possible to try estimating the desired WTPs using aggregated data and estimation techniques appropriate to such data. Alternatively, a narrower version of the model could be estimated using disaggregated data, such as a route choice model where the observed data includes known bottlenecks (e.g. Auckland Harbour Bridge, or Hutt Motorway).[19]

We recommend attempting to supplement the HTS data, despite it not being already sufficient for DCA, because it represents an important record of actual travel decisions in New Zealand. This means it is currently the only comprehensive New Zealand dataset containing information on actual transport user choices that might be used to estimate their preferences for different transport attributes. As such, if it can be successfully augmented as recommended, it should allow estimation of transport users' actual preferences, free of hypothetical biases.

1.5 Report Structure

The balance of this report is structured as follows:

- Section 2 backgrounds the DCA approach, explaining how it can be used to estimate transport users' preferences for different transport attributes;
- Section 3 summarises the HTS data, with particular focus on how it might be used for DCA (since it was not created specifically for that purpose);
- Section 4 discusses how the HTS is either suitable, or requires supplementary data, to measure variables of interest for DCA;
- Section 5 refers to how estimates of WTPs for transport attributes are used in transport analyses of relevance to the GPS output classes; and
- Section 6 sets out our recommended DCA approach, and concludes.

[19] Such bottlenecks naturally define routes, even if the existing HTS data does not include explicit route data.

2 Estimating Transport User Trade-Offs using Discrete Choice Analysis – Some Background

2.1 Outline of the Approach

This section provides some introductory background on the DCA methodology, with particular reference to estimating traveller's tastes for different attributes of transport modes. It begins by introducing the basic framework for DCA, namely consumers choosing an alternative from among the range of alternatives available to them so as to maximise their satisfaction. It then explains how consumers' tastes for different attributes of travel modes – such as their tastes for travel time, travel time reliability, ride quality and safety – are represented, interpreted and estimated. The simplest case in which all travellers have the same tastes for each attribute is presented first. Then, the more relevant case – in which different travellers can have different tastes for each attribute – is discussed. Finally, the application of different data types in DCA models, and associated estimation techniques, are briefly mentioned.

2.2 Travellers Choose the Transport Mode Giving them the Greatest Satisfaction

DCA is founded on the random utility model of consumer choice. It supposes that consumers (e.g. travellers) will choose that alternative (e.g. travel mode), from among a list of alternatives, that gives them the greatest satisfaction, or "utility". For example, a traveller that has the choice between walking to work, or taking a bus, will choose the transport mode that gives them the most utility. That utility will depend on the attributes of each mode. So, for example, walking provides exercise and has no direct financial cost, but is slow and involves the risk of encountering bad weather, or being run over by a car. Conversely, taking a bus provides no exercise, requires sharing space with strangers, and has a direct financial cost. But it also requires less travel time, enables other activities such as reading, and little risk from bad weather or car crashes.

2.3 How Travellers Choose Transport Mode, Supposing they have the same Preferences

Suppose, for now, that each traveller has the same preference, or "taste", for each transport mode attribute. Their utility can be represented as the sum of two parts – one not random (V_j), and the other random (ϵ_{ij}). Specifically, traveller i choosing transport mode j is assumed to derive utility U_{ij} having the following form:

$$U_{ij} = V_j + \epsilon_{ij}$$

A common interpretation of the random term ϵ_{ij} is that it captures factors influencing traveller i's utility from transport mode j where those factors are not observable by the researcher. For example, it could capture whether the traveller enjoys the view along the walking route, or if the bus driver is polite.

In turn, the non-random – i.e. observable – utility component V_j is typically represented in linear form:

$$V_j = \alpha_j + \beta_{j1}x_{j1} + \ldots + \beta_{jn}x_{jn}$$

where α_j is an "alternative-specific constant", measuring the average utility provided by mode j.[20] Here $x_{j1}, \ldots, x_{jn}$ represent n different observable attributes of transport mode j (including travel time, direct cost, ability to read, etc), while $\beta_{j1}, \ldots, \beta_{jn}$ represent unknown travellers' tastes for each such attribute of that mode. For example, the taste parameter on travel time or direct cost would be expected to be negative, indicating that travellers dislike travel time and direct travel cost. Conversely, it would be positive if travellers value being able to read while travelling. Each taste parameter could differ by mode – i.e. the dislike of travel time might be higher or lower for walking versus taking a bus.

The objective of DCA is to use observations on travellers' actual (or stated) transport mode choices, plus information about the attributes of each transport mode, to statistically estimate the unknown taste parameters $\beta_{j1}, \ldots, \beta_{jn}$. To do this, assumptions are required on the statistical properties of the random components of utility, ϵ_{ij}.

2.4 Simplest DCA Model – Multinomial Logit

The simplest form of DCA is the so-called multinomial logit (MNL), or simple logit, model. In this model the random component is assumed to be independently and identically-distributed with a Type I Extreme Value statistical distribution.[21] With this assumption, it can be shown that the probability that traveller i chooses transport mode j from among the J available trans-

[20] More precisely, this measures average utility *relative to* some reference choice alternative (i.e. mode).

[21] The reason for this particular assumption is that it enables simple expressions for traveller's mode choice probabilities to be derived.

port modes has the simple form:[22]

$$P_j = \frac{exp\,(V_j)}{\sum_{k=1}^{J} exp\,(V_k)}$$

Since the non-random utility components V_j depend on the taste parameters $\beta_{j1}, \ldots \beta_{jn}$, expressions such as this enable estimation of those parameters.[23]

2.5 Interpreting Preferences, and Willingness to Pay

Because of the random utility formulation assumed above, the taste parameters show how a traveller's utility changes for small changes in transport mode attributes. For example, the taste parameter for travel time shows how travellers' utility falls as travel time increases. Such taste parameters measure travellers' "marginal utilities" for each attribute. As such, the ratio of any two taste parameters measures how travellers trade off one transport mode attribute against another. For example, the ratio of travel time and "protection from the weather" attributes measure how travellers trade off travel time against protection from the weather.[24]

In economic terms, the ratio of any two taste parameters measures travellers' "marginal rate of substitution" between the two attributes in question. While this is useful in conceptual terms, a more practical measure of how travellers trade off the attributes of a transport mode reduces the trade-off to monetary values. This is easily achieved if at least one of the transport mode attributes is a financial variable (e.g. direct travel cost). For example, the (negative of the) ratio of travellers' taste parameter for travel time to their taste parameter for travel cost shows how they trade off travel time for money. In other words, it shows their WTP for travel time – i.e. VOT.[25]Thus, for example, if traveller's taste for travel time was estimated to be $\hat{\beta}_{Time} = -5$,

[22]The probability that traveller i chooses transport mode j from among the J available transport modes is calculated as the probability that mode j gives the highest utility U_{ij} from among all the alternatives. See Train (2009) for details.

[23]More precisely, the choice probabilities, which depend on the taste parameters, can be used to derive a log-likelihood function. The parameter estimates are then chosen to maximise that log-likelihood.

[24]More precisely, holding the travellers' total utility constant. Also, a negative sign is introduced in such calculations, due to how WTP is defined and computed. Train (2009, p. 39) provides details.

[25]Note that how travellers trade off transport mode attributes can be computed directly from ratios of taste parameters. It can also be computed indirectly using ratios of WTP for each attribute (since each WTP has a common denominator – the taste parameter for travel time).

this means their utility falls by 5 units for every unit increase in travel time. If their taste for direct travel cost was estimated to be $\hat{\beta}_{Cost} = -2$, then their utility falls by two units for every unit increase in direct travel cost. Using these hypothetical figures, travellers' WTP for travel time savings – their VOT – is estimated as:

$$VOT = WTP_{\text{Travel Time Savings}} = -\frac{\hat{\beta}_{Time}}{\hat{\beta}_{Cost}} = -2.5$$

This implies that travellers are prepared to incur an extra 2.5 units of travel cost in order to save one additional unit of travel time, holding their overall utility constant. For example, if travel cost is in dollars, and travel time in lots of 5 minutes, then the above VOT indicates travellers would pay \$2.50 to save 5 minutes in travel time.[26]

2.6 Allowing for Travellers to have Different Tastes – Mixed Logit Model

The simple DCA formulation above has several known limitations.[27] These arise due to the restrictive assumption that the random component in a traveller's utility is independently and identically distributed. Furthermore, transport studies commonly find that tastes for different transport mode attributes vary across travellers. Hence, for example, retired people have a different dislike of travel time than working people or children. As a consequence, the standard approach for estimating DCA models in transport applications is the so-called Mixed Logit (ML), or random-coefficient logit, model.

An important feature of the ML model is that it does not require taste parameters to be the same across travellers. Hence, instead of assuming that $\beta_{j1}, \ldots, \beta_{jn}$ are the same across all travellers, traveller i's utility from choosing transport mode j now writes with a form such as:

$$U_{ij} = \alpha_{ij} + \beta_{ij1}x_{j1} + \ldots + \beta_{ijn}x_{jn} + \epsilon_{ij}$$

[26]If the taste parameter in the numerator was for a desirable travel mode attribute such as ride quality, then the resulting WTP – in this case VOQ – would be a positive number. This indicates that travellers would be prepared to pay more direct travel cost if they also enjoyed higher ride quality. This is intuitive, since travel cost reduces utility, while ride quality increases it. Conversely, in VOT, both travel time and travel cost are likely to reduce utility, in which case one of them has to fall if the other rises for utility to be left unchanged.

[27]See Train (2009), Small and Verhoef (2007).

Notice that each taste parameter is now specific to the traveller.[28] Furthermore, each taste parameter is now allowed to vary across individuals for both observable and unobservable (to the researcher) reasons. For example, each taste parameter is assumed to now depend on characteristics (i.e. demographics) of each traveller, such as income levels, gender, age, work status, etc. This introduces observable variation into the estimation of taste parameters.

Furthermore, it is well documented that there is a large element of unobserved variation explaining differences in traveller's taste coefficients.[29] This is captured by assuming that taste coefficients also include unobserved, random components. By making assumptions about the statistical properties of these unobserved components, it is possible to derive an expression for the probability that traveller i (with particular demographics) chooses transport mode j (with particular attributes) from among the J available mode alternatives. As for the simpler MNL model, this expression depends on the unknown taste parameters $\beta_{ij1}, \ldots, \beta_{ijn}$. However, unlike the MNL model, the ML model produces an expression for P_j that cannot be solved directly.

Once again, given expressions for each choice probability, it is possible to use statistical techniques to estimate the required taste parameters. In the ML case, however, this means a *distribution* of taste parameters is produced, instead of single *point estimates*, reflecting how tastes differ across travellers. In turn, this means that ratios of taste parameters – including WTP estimates – are also distributions instead of point estimates. In short, the ML model provides information about how different kinds of travellers differ in how they trade off different transport mode attributes.

2.7 Data Types and Estimation

As mentioned in Section 1, DCA models can be applied using data on travellers' actual mode choices (i.e. revealed-preference, or RP, data). Alternatively, they can be applied using data on which modes travellers state they would choose in stated preference choice experiments (i.e. stated preference, or SP, data). Finally, they can also be applied using combinations of RP and SP data.

Whether the relevant data is RP, SP or both, there is a fundamental distinction between data types that is relevant for how DCA models are applied. Specifically, DCA models can be applied using:

[28] In practice, only some of the taste parameters might be made varying across travellers, with others assumed constant across travellers. This can be for reasons of tractability, or on theoretical grounds.

[29] E.g., see Small (2012).

- Disaggregated data – i.e. data at the level of individual travellers' mode choices; or

- Aggregated level – i.e. data on travellers' mode choices for groups of travellers (e.g. by locale).

In the former case, it is necessary to have data not only on which mode each traveller chose, but also on modes they might have chosen but did not. In the latter, simply knowing the mode shares of each mode actually chosen is adequate. In either case it is necessary to have information on the attributes of all modes (whether chosen or non-chosen).

Disaggregated data is the most desirable for estimating models such as ML models, given the need to distinguish different types of travellers in order to estimate how their tastes for transport mode attributes vary. It also enables relatively simple estimation techniques to be used. Indeed, even the ML model, which is more complicated than the MNL model, can be estimated using conventional techniques.[30]

If quality disaggregated data is not available, then provided there is sufficient variation in aggregated data, this too can be used to estimate even demanding models like ML models. In this case, however, much more sophisticated estimation algorithms are required.[31] Furthermore, results can be very sensitive to how they are estimated, requiring careful attention to ensuring robustness of results.[32]

[30]Specifically, simulated maximum likelihood estimation – see Train (2009).

[31]The approach was pioneered by Berry et al. (1995), with transport applications discussed in Small and Verhoef (2007). It has been much elucidated, such as by Nevo (2000), with Girotti and Meade (2015) providing a summary of both model specification and estimation approach.

[32]Knittel and Metaxoglou (2014).

3 HTS Description

3.1 Overview

The HTS is an ongoing survey that details people's day-to-day travel in New Zealand. For each surveyed household, it collects information on all trips taken, over two consecutive surveyed days, including the origin and destination, departure and arrival times, modes of travel, and purpose. This is the primary source of data on travel behavior in New Zealand. Its wide coverage with respect to geographic and time dimensions makes it particularly useful for a broad investigation on travel patterns.

Detailed description of the HTS, including survey methods and actual questionnaires used, is available at many other sources, including the MoT website.[33] To avoid duplication, the description in this report will only focus on aspects of the data that pertain to our proposed DCA. One important aspect that is emphasized here is the investigation on the sparseness of observations along various categorization, along dimensions such as geography, time, and mode. This gives guidance to what categories and divisions will actually be feasible in our DCA using the HTS. Some of these were never explored explicitly in previous studies that used the HTS. Another aspect we are constantly mindful of is the appropriateness of survey weights present in the HTS and its data coverage. These will ultimately dictate how far we can generalize our estimation results from the HTS to all trips, taken on all modes, by all people in New Zealand.

The HTS is an ongoing, continuous survey. However, a new survey method using GPS units and online forms was introduced in October 2015. This demarcates the "current survey" from the "old survey", which used paper "memory joggers" and paper questionnaires. At the time of writing of this report, data from the current survey using GPS units are not yet available. This report relies on the "old survey" only and will limit its description to that. The "old survey" data we have spans from survey year 2003/04 to survey year 2013/14.[34]

Also, this section describes the parts of the HTS that was supplied to us for this report only. It is our understanding that a few smaller data sets were omitted among the files

[33] `http://www.transport.govt.nz/research/travelsurvey/`

[34] The survey year does not coincide with the calendar year. For example, the survey year 2003/04 starts in July 2003.

shared with us, for example, the accident data and alcohol data.

3.2 Sampling Method and its Consequences

The HTS employs a stratified sampling method.[35] The strata first divides New Zealand into 14 regions. Most of these regions contain three area types: main urban, secondary urban, and rural.[36] Within each region, meshblocks are sampled proportional to size, without replacement. Within each meshblock sampled, one out of eight households is sampled in random order. In a surveyed household, all people within that household are surveyed. Finally, each surveyed household is assigned two consecutive travel days. All trips taken on those two days, by all people in that household, are recorded.

Because of this specific stratified sampling method, the households and trips surveyed are far from a random sample. There is clustering on many levels, largely to minimize survey cost. The trips surveyed are *not* a random sample out of all trips taken in that year in New Zealand; instead they are clustered by surveyed days and persons. The persons surveyed are *not* a random sample out of all persons in New Zealand; instead they are clustered by surveyed households. The households surveyed are *not* a random sample out of all households in New Zealand; instead they are clustered by meshblocks. As a result, each each trip, person, and household in the HTS has a sampling weight, which is used in all estimation. In addition, there is a non-response weight and post-stratification weight[37] for households.

We are mindful of the consequences of this stratified sampling and resultant clustering. Firstly, in terms of econometric estimation, this means that standard errors will have to be adjusted in order to be reliable, because a clustered sample is generally *less* varied than a truly random sample. If the standard error were uncorrected for clustering, it would be smaller than the "true" standard error, thus falsely inflating statistical significance in hypothesis tests. Secondly, in terms of geographic coverage, the HTS spans all 14 regions in New Zealand, but only a very small number of meshblocks are sampled in

[35] This subsection benefits greatly from Michael Keall's presentation slides, titled "Travel Survey Design and Estimation Issues", supplied by the ministry.

[36] Four regions (Gisborne, Hawke's Bay, Northland, and West Coast) are exceptions to this. These exceptions are clearly visible in tables to follow.

[37] To adjust for non-coverage in survey.

each region. In other words, the vast majority of meshblocks in New Zealand have no presence in the HTS. Whether the results from the HTS is generalizable to the whole of New Zealand depends on the degree of similarity between sampled meshblocks and non-sampled ones. Thirdly, in terms of longitudinality, the HTS has no repeated observation of any person or household.[38] In statistical language, the HTS is not a panel data set. (Strictly speaking, it is *not* a cross-section data set either, since households are sampled continuously throughout the year, instead of all at once at the same point in time.) The lack of repeated observations within a person or a household means that we cannot disentangle any person- or household-specific effects from the observed data. The lack of a cross-sectional property means that, if the effect of time (time of day; day of week; month of year) on travel behavior is not uniform across all persons/households, we cannot disentangle the effect of time from the effect of the specific person/household.

3.3 Data Type and Variables

The data sets used for this report include the Address data, Household data, Person data, Trip data, and Vehicle data. Here we list the key features and variables relevant to our DCA in each data set, in order.

The Address data contains address information on all locations mentioned in sampled trips. Each address contains a description,[39] a physical address or a street intersection, and geo-coded x- and y-coordinates. Another variable indicates the level of geo-code accuracy. This can be merged into the Trip data.

The Household data contains the two consecutive survey days, home structure type,[40] household type,[41] number of residents and eligible persons, response status, vehicle and bicycle ownership, meshblock, area type, and region.

The Person data contains demographic variables such as sex, age, relationship to head of household, employment status indicators,[42] school or work address number (to be

[38] A very few number of households were sampled a second time in a different survey year in the HTS, but this exception is too small to be of use.

[39] E.g., "home", "school", "bank".

[40] E.g., "separate house"; "2 flats or houses joined together", etc.

[41] E.g., "person living alone"; "married/de facto couple only", etc.

[42] E.g., "student full time"; "student part time"; "work full time"; "work part time"; "keeping house", etc.

merged with the Address data), occupation code, personal income range. There are variables on historic travel behavior, such as cycling in previous year/previous four weeks, use of public transport in previous month, lifetime driving experience, kilometres driven in previous year, and driver licence status and type.

The Trip data contains the person number, trip day, order of this leg in overall journey, overall journey purpose, origin and destination address numbers, start and arrival times, travel mode, and travel distance. Trip distance is further broken down into road type, such as state highways, open roads, and urban roads.

The Vehicle data contains the household number, vehicle manufacturer, model, year, type,[43] engine capacity, owner, and fuel used.

3.4 Data Coverage

In this subsection we explore the sparseness of data along various dimensions suggested above, using tables. These tables are indicative of the level of richness of the HTS data, which will dictate how rich any DCA model could be. Whenever a survey response variable is considered, only households with "full responses" are used, and this is clearly stated in the footnotes of all tables involved.[44] All tables in this section make use of all 11 years ("years 2-12" in HTS) of data available; this is also clearly stated in all table titles. (Other researchers interested in changes in travel behavior *across* these years might split the variables of interest along the year dimension. This would also be legitimate for a DCA study. However, there are other dimensions likely to be more important for a DCA on travel mode choice. And as these tables will show, many cells are already fairly sparse such that further slicing will lead to unreliable estimation results.) Lastly, since one of the main purposes of these tables is exploration on data sparseness, all tables show observation counts instead of row or column percentages, and no weights are used.

Table 1 shows the total number of meshblocks in each region, and those present in the HTS. The percentage of meshblocks sampled is generally below 4%. In addition, the HTS is designed to under-sample the largest cities (Auckland and Wellington) and over-sample

[43] E.g., "car/station wagon"; "van/ute/passenger van"; "4 wheel drive", etc.

[44] In terms of "full responses", this is the strongest filter in the HTS. All tables in this subsection use the same filter for consistency and comparability. For analysis on lower levels, such as the person-level, "persons with full responses" is a weaker filter that yields more usable observations.

Table 1: Number of Meshblocks per Region: Total vs. HTS, years 2-12

Region	Total	Covered in HTS	Percentage
Area Outside	23	–	–
Auckland	11768	209	1.78%
Bay of Plenty	2860	61	2.13%
Canterbury	5915	215	3.63%
Gisborne	533	37	6.94%
Hawke's Bay	1824	46	2.52%
Manawatu-Wanganui	3113	60	1.93%
Nels-Marlb-Tas	1555	46	2.96%
Northland	2047	55	2.65%
Otago	2747	73	2.66%
Southland	1739	70	4.03%
Taranaki	1557	52	3.34%
Waikato	5164	90	1.74%
Wellington	5237	93	1.78%
West Coast	515	43	8.35%

the smallest regions. Thus, while the HTS has breadth, it does not have nearly enough depth for analysing rich local variations in travel behavior, such as commuter mode choice to Auckland CBD across suburbs. It should also be noted that the the same meshblocks might not be present in all survey years. In year 5, new meshblocks were rotated into the sample to replace ones that were almost exhausted.

One dimension of time that would be desirable to capture is seasonality—although we are unlikely to be able to capture its changes across years. We deem the former more important than the latter because monthly seasonality has much more persistence (largely due to the school calendar), which makes it an indispensable component in travel behaviour forecasting. As Table 2 shows, the continuous nature of the HTS gives it good coverage in all twelve months of the year, although it is not perfectly even.

Table 3 tabulates the same set of household observations by day-of-week. Again, all cells are reasonably covered, although not uniformly across the row: in many regions the middle of the week contains more observations. The use of day-of-week is not limited to a comparison between weekdays and weekends. (On the matter of distinguishing between work and non-work travels, the "journey purpose" variable is likely more reliable, since not all trips on a workday are work trips, and not all work trips happen on workdays.)

Table 2: Number of Households with Full Response by Region–Month in HTS, years 2-12

Region	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Auckland	136	148	175	130	152	175	141	148	157	128	167	141
Bay of Plenty	59	31	56	68	39	47	57	53	36	61	56	46
Canterbury	181	182	216	261	217	204	116	171	235	204	175	164
Gisborne	14	25	22	16	16	31	7	21	23	12	31	31
Hawke's Bay	27	29	31	33	26	34	30	28	37	33	25	28
Manawatu-Wanganui	41	29	15	37	46	22	40	39	9	43	33	30
Nels-Marlb-Tas	45	26	28	42	47	25	23	36	14	41	38	26
Northland	18	22	33	46	27	54	27	26	26	32	39	26
Otago	37	37	48	50	36	44	52	46	45	62	49	42
Southland	28	24	40	23	17	24	38	27	35	36	28	38
Taranaki	27	36	41	22	33	31	32	31	44	42	20	51
Waikato	79	66	69	80	79	68	83	64	64	86	74	59
Wellington	69	60	69	84	58	86	66	68	63	70	46	59
West Coast	25	18	15	20	18	20	34	17	9	26	13	18

*Note: The sum of each row is **not** the total number of households with full response in that region, since each household is sampled for two consecutive days, and in some cases these two days span two consecutive months.*

Table 3: Number of Households with Full Response by Region–Day-of-Week in HTS, years 2-12

Region	Mon	Tues	Wed	Thurs	Fri	Sat	Sun
Auckland	499	480	483	487	515	559	525
Bay of Plenty	145	145	160	199	206	165	156
Canterbury	680	680	640	631	615	627	663
Gisborne	62	75	76	74	78	67	52
Hawke's Bay	83	103	116	112	95	97	98
Manawatu-Wanganui	119	110	99	103	109	102	110
Nels-Marlb-Tas	84	88	92	125	141	118	92
Northland	131	110	98	107	100	83	101
Otago	135	158	172	161	147	154	139
Southland	105	122	107	86	89	91	96
Taranaki	106	105	118	107	99	119	128
Waikato	228	232	258	275	267	229	215
Wellington	217	234	245	247	235	204	194
West Coast	50	58	63	64	71	83	59

*Note: The sum of each row is **not** the total number of households with full response in that region, since each household is sampled for two consecutive days.*

This is also useful for incorporating external data into a DCA, such as road traffic volume or public transport frequency by day of week, both of which have strong persistence by day-of-week.

Our exploration now turns to survey response variables, such as trips taken and their modes and purposes. All tables from this point onward are tabulated by the 14 regions and three area types (main urban area; secondary urban area; rural). This is the level of geographic detail we ideally wish to use in a DCA. Table 4 shows the number of observed trips by all available travel modes in the HTS. Not surprisingly, some modes are very sparsely populated. In regions other than Auckland and Wellington, "train" and "ferry" should probably be combined with "bus" as one "public transport" mode. "Taxi" and "mobility scooter" should probably be combined with "other". Trips taken by "plane" could either be combined with "other" or discarded, depending on whether the analysis focuses on relatively short local trips, or inter-city trips. This table also shows that Gisborne, Hawke's Bay and Northland have no secondary urban areas, while West Coast has no main urban area. Table 5 shows the same data, but at higher level of aggregation.

Table 4: Total Number of Trips by Mode, Region, Urban/Rural in HTS, years 2-12

Region	Area	Walk	Vehicle Driver	Vehicle Passenger	Bicycle	Train	Bus	Ferry	Plane	Taxi	Other	Mobility Scooter
	MUA	13556	46908	25288	444	382	2268	216	102	243	408	11
Auckland	SUA	317	1278	632	20	8	19	0	5	13	3	0
	Rural	337	1720	922	10	4	58	2	4	1	14	0
	MUA	3089	12532	6379	420	0	342	2	17	40	35	31
Bay of Plenty	SUA	275	1066	738	43	0	66	0	2	5	1	0
	Rural	379	2433	1328	61	1	139	0	4	8	4	0
	MUA	12689	33689	16868	1785	4	2125	4	122	262	91	47
Canterbury	SUA	1168	4677	1766	142	0	141	0	6	11	10	0
	Rural	5642	16490	8600	725	0	584	3	61	44	24	22
	MUA	1742	5196	3140	213	3	59	3	13	18	18	19
Gisborne	SUA	0	0	0	0	0	0	0	0	0	0	0
	Rural	362	1806	960	51	0	76	0	10	1	13	0
	MUA	1869	7911	4183	303	0	276	2	5	40	1	18
Hawke's Bay	SUA	0	0	0	0	0	0	0	0	0	0	0
	Rural	367	1844	910	30	0	90	0	1	2	5	0
	MUA	1109	5209	2282	275	0	106	8	11	35	27	17
Manawatu-Wanganui	SUA	472	3148	1544	69	4	34	2	4	22	6	0
	Rural	762	3774	1908	40	3	127	0	1	8	17	4
	MUA	1879	5933	2941	316	0	67	0	16	38	37	9
Nels-Marlb-Tas	SUA	817	2754	1196	132	2	19	0	6	11	7	18
	Rural	972	3568	1482	148	7	148	8	6	1	30	2
	MUA	1553	3947	2250	48	0	97	2	7	14	3	0
Northland	SUA	0	0	0	0	0	0	0	0	0	0	0
	Rural	1562	6307	3412	51	0	332	11	9	9	19	0
	MUA	4592	10422	5264	261	7	334	0	34	80	33	27
Otago	SUA	352	1364	637	58	0	14	0	2	7	0	0
	Rural	1489	4158	1491	121	0	122	1	14	19	19	0
	MUA	1078	5214	2607	133	0	146	0	9	33	25	8
Southland	SUA	295	1187	546	31	0	10	0	0	2	2	8
	Rural	513	4012	1818	65	0	319	2	9	7	11	4
	MUA	1458	5515	2701	138	0	79	5	12	26	3	16
Taranaki	SUA	510	1756	784	85	0	22	0	0	5	5	0
	Rural	806	3939	2098	52	6	111	4	5	4	15	3
	MUA	3202	14580	7417	370	14	371	6	20	59	53	12
Waikato	SUA	595	2250	1310	81	9	41	0	1	11	12	13
	Rural	1860	9052	4318	164	10	277	3	13	21	54	10
	MUA	11590	22558	12691	426	675	1608	33	106	211	41	10
Wellington	SUA	319	1519	616	45	14	9	1	3	8	0	0
	Rural	669	2097	818	60	43	45	4	3	0	7	13
	MUA	0	0	0	0	0	0	0	0	0	0	0
West Coast	SUA	566	2159	1016	67	0	16	0	1	12	3	0
	Rural	1351	4553	2372	225	0	104	2	8	37	25	3

Note: Only trips from households with full responses are included.

Table 5: Total Number of Trips by Mode, Region, Urban/Rural in HTS, years 2-12

Island	Area	Walk	Vehicle Driver	Vehicle Passenger	Bicycle	Train	Bus	Ferry	Plane	Taxi	Other	Mobility Scooter
North Island	MUA	39168	124356	66331	2637	1074	5206	277	293	686	589	134
	SUA	2488	11017	5624	343	35	191	3	15	64	27	13
	Rural	7104	32972	16674	519	67	1255	24	50	54	148	30
South Island	MUA	20238	55258	27680	2495	11	2672	4	181	413	186	91
	SUA	3198	12141	5161	430	2	200	0	15	43	22	26
	Rural	9967	32781	15763	1284	7	1277	16	98	108	109	31

Note: Only trips from households with full responses are included.

Trips are chained together to form a journey in the HTS.[45] For many variables that follow, "journey" is a more natural level of observation than "trip". Table 6 tabulates overall journey purpose. The HTS allows for 15 different journey purposes, including "home", three types of "work" journeys, "education", "shopping", and other personal, social, or leisure purposes. The full list of purpose codes are listed in the footnote of Table 6.[46] Journeys whose purpose is to "leave country" (code 14) can be discarded in a DCA on local travel. Journeys with "social welfare" purpose (code 7) are quite sparse, and could be combined with "personal business / service" (code 8). All other purposes are reasonably well populated.

[45] A "trip" is considered a unit of travel with a single mode. For example, when a person walks 15 minutes to the bus stop, then takes a 20 minutes bus ride, this is considered two separate trips. If the time gap between these two trips is less than 30 minutes, and that one of these trips has no dedicated purpose of its own (here, the purpose of the 15 minutes walking trip is to "change mode"), these trips are combined into one single journey.

[46] There is no journey purpose code "12" because this value is assigned to trip purpose "change mode". By definition of a journey, this is purpose code is eliminated when trips are aggregated to journeys.

Table 6: Total Number of Journeys by Purpose, Region, Urban/Rural in HTS, years 2-12

		Purpose Code														
Region	Area	1	2	3	4	5	6	7	8	9	10	11	13	14	15	16
Auckland	MUA	30513	7108	190	3485	3721	10300	94	3694	597	7998	5144	11275	17	8	548
	SUA	788	195	0	68	78	374	2	131	26	206	157	207	3	0	7
	Rural	969	183	16	232	96	381	0	124	33	367	159	384	0	1	27
Bay Of Plenty	MUA	7760	1563	61	983	672	2972	21	1068	245	2434	1641	2532	2	2	243
	SUA	704	166	15	108	95	260	2	81	18	264	104	307	2	0	12
	Rural	1436	260	25	250	180	469	5	264	37	455	209	553	0	0	42
Canterbury	MUA	22017	5021	142	2437	2370	8886	46	2528	430	8485	3247	6842	13	5	594
	SUA	2552	716	29	358	200	988	1	368	69	937	485	744	1	0	74
	Rural	10570	2409	90	1487	1185	3836	21	1360	231	3974	1642	3570	3	4	291
Gisborne	MUA	3287	758	30	371	348	1204	7	579	89	1206	542	1360	0	0	101
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	984	218	8	224	142	364	3	272	26	375	139	350	3	0	40
Hawke's Bay	MUA	4845	1055	39	577	471	2118	20	582	80	1673	721	1843	0	0	99
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	1024	225	14	218	137	376	1	183	33	368	124	426	0	1	28
Manawatu-Wanganui	MUA	3074	818	66	365	337	1141	14	410	71	975	507	945	1	0	76
	SUA	1779	508	13	290	145	669	5	207	66	588	234	662	0	0	31
	Rural	2114	466	24	400	226	836	2	288	51	916	368	740	1	0	40
Nels-Marlb-Tas	MUA	3633	793	26	560	379	1400	7	495	99	1199	786	1335	0	0	84
	SUA	1575	481	49	265	111	570	1	243	41	592	318	459	0	0	46
	Rural	1944	446	38	459	196	731	2	322	40	664	447	712	1	0	51
Northland	MUA	2226	488	19	410	208	1039	11	321	58	927	423	1049	0	0	80
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	3584	790	44	658	395	1460	10	519	98	1467	643	1314	0	0	114
Otago	MUA	6286	1603	64	916	659	2683	17	1071	160	2075	1460	2526	1	3	202
	SUA	830	205	7	130	74	247	1	132	32	216	155	336	0	0	10
	Rural	2449	808	44	541	199	775	7	367	43	789	461	672	4	5	28
Southland	MUA	3023	763	58	363	374	1118	5	485	76	759	671	1252	1	7	120
	SUA	738	151	9	60	52	348	2	92	18	219	198	183	0	0	4
	Rural	2201	471	28	442	323	828	8	358	64	531	499	697	0	0	78
Taranaki	MUA	3130	828	24	465	298	1219	13	530	118	1015	686	1203	1	0	149
	SUA	1061	219	48	163	80	409	1	219	22	347	163	325	0	0	29
	Rural	2272	408	37	441	190	801	11	441	46	821	404	900	0	0	78
Waikato	MUA	8228	2109	71	1347	832	3104	49	1302	254	2789	1518	3210	5	0	280
	SUA	1428	349	3	164	180	505	6	227	26	417	250	606	0	0	26
	Rural	5041	1285	48	963	584	1788	28	894	135	1601	911	1855	1	2	145
Wellington	MUA	14352	4108	137	2059	1336	6118	26	1826	319	5205	2814	5767	8	1	281
	SUA	854	226	18	178	55	298	0	83	31	313	137	238	0	0	10
	Rural	1166	258	26	255	71	486	2	179	37	400	249	362	0	0	11
West Coast	MUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	SUA	1264	342	9	243	78	445	2	318	24	361	316	291	0	0	26
	Rural	2992	716	43	366	199	1024	9	625	54	839	738	769	0	0	89

Note: Only journeys from households with full responses are included. Purpose codes: 1 = Home; 2 = Work (main job); 3 = Work (other job); 4 = Work (employer's business); 5 = Education; 6 = Shopping; 7 = Social welfare; 8 = Personal business / services; 9 = Medical / dental; 10 = Social visit / entertainment; 11 = Recreational; 13 = Accompany someone else; 14 = Left country; 15 = Other; 16 = Overnight lodging.

The next two tables, Tables 7 and 8, present the distributions of journey distances (in kilometres) and duration (in minutes) using percentiles. At the tenth percentiles, journeys are shorter than half a kilometer and last for less than five minutes in all regions. Journeys are not that long even at the ninetieth percentile, at about 30-40km and rarely lasting more than 40 minutes. This shows that the vast majority of journeys covered by the HTS are relatively short, local, day-to-day journeys. Indeed, the HTS would be unsuited for modelling inter-city or inter-island travel.

Another important time dimension is journey departure hour. Similar to day-of-week, departure hour is particularly useful for incorporating external data such as road traffic volume or public transport frequency, both of which have substantial variation throughout the day. Departure hour is explored in Table 9 (split horizontally into three pages due to its width). The twenty-four hours of the day are divided into forty-eight half-hour slots, each represented by a column. At this level of disaggregation, many cells contain very few observations, except in Auckland, Canterbury, and Wellington MUA's. In all other regions, the early morning hours before 5:00am or so should probably be aggregated together.

Table 7: Percentiles of Journey Distance (km) by Region, Urban/Rural in HTS, years 2-12

		Percentiles				
Region	Area	10th	30th	50th	70th	90th
Auckland	MUA	0.700	1.933	3.774	7.571	18.89
	SUA	0.476	1.387	2.343	8.303	35.63
	Rural	0.410	2.314	6.280	12.12	39.45
Bay Of Plenty	MUA	0.606	1.822	3.348	5.670	13.63
	SUA	0.616	2.070	3.414	5.120	20
	Rural	0.537	2.203	4.477	10.60	29.07
Canterbury	MUA	0.641	1.762	3.394	6.158	13.31
	SUA	0.521	1.624	2.673	4.074	15.69
	Rural	0.480	1.214	2.971	12.35	27.11
Gisborne	MUA	0.477	1.382	2.330	3.640	8.377
	SUA	.	.	.	.	.
	Rural	0.442	1.509	4.202	12.09	25.74
Hawke's Bay	MUA	0.523	1.617	3.087	6.460	16.21
	SUA	.	.	.	.	.
	Rural	0.326	0.959	2.321	9.440	38.25
Manawatu-Wanganui	MUA	0.631	1.735	2.808	4.098	8.200
	SUA	0.619	1.578	2.630	5.045	21.82
	Rural	0.496	1.690	4.358	10.56	30.23
Nels-Marlb-Tas	MUA	0.496	1.342	2.634	5.455	11.25
	SUA	0.446	1.250	2.143	3.750	10.06
	Rural	0.360	1.210	3.501	10.44	28.03
Northland	MUA	0.442	1.500	3.369	6.196	13.32
	SUA	.	.	.	.	.
	Rural	0.375	1.500	4.473	10.03	28.87
Otago	MUA	0.450	1.315	2.553	4.960	13.35
	SUA	0.460	1.339	2.138	4.488	8.898
	Rural	0.380	1	1.939	7.102	20.90
Southland	MUA	0.603	1.680	2.754	4.291	10.51
	SUA	0.412	0.782	1.150	1.799	18.56
	Rural	0.590	2.140	5.850	14.54	33.37
Taranaki	MUA	0.516	1.440	2.626	4.354	9.405
	SUA	0.383	1	1.789	3.853	17.14
	Rural	0.440	1.550	4.952	15	30.75
Waikato	MUA	0.633	1.727	3.621	6.660	19.80
	SUA	0.528	1.250	2.029	5	28.98
	Rural	0.486	1.472	3.710	11.28	33.80
Wellington	MUA	0.450	1.505	3.253	6.951	21.60
	SUA	0.508	1.103	1.770	2.748	11.18
	Rural	0.429	1.300	2.595	6.554	26.46
West Coast	MUA	.	.	.	.	.
	SUA	0.360	1.313	2.290	3.792	10.29
	Rural	0.307	0.907	1.678	4.255	25

Note: Only journeys from households with full responses are included.

Table 8: Percentiles of Journey Duration (Minutes) by Region, Urban/Rural in HTS, years 2-12

		Percentiles					
Region	Area	10th	30th	50th	70th	90th	*N*
Auckland	MUA	5	10	15	20	35	84675
	SUA	5	5	10	20	40	2242
	Rural	4	8	13	20	45	2972
Bay Of Plenty	MUA	5	6	10	15	30	22189
	SUA	4	5	10	15	30	2138
	Rural	4	5	10	20	39	4185
Canterbury	MUA	5	8	11	17	30	63044
	SUA	4	5	9	10	25	7520
	Rural	3	5	10	20	35	30649
Gisborne	MUA	5	5	10	10	20	9882
	SUA	.	.	.	.	.	0
	Rural	3	5	10	15	35	3148
Hawke's Bay	MUA	4	5	10	15	29	14123
	SUA	.	.	.	.	.	0
	Rural	2	5	10	15	40	3158
Manawatu-Wanganui	MUA	4	5	10	10	21	8799
	SUA	3	5	7	15	30	5197
	Rural	3	5	10	15	35	6463
Nels-Marlb-Tas	MUA	4	5	10	15	25	10793
	SUA	4	5	8	10	20	4751
	Rural	3	5	10	15	35	6053
Northland	MUA	3	5	10	15	30	7259
	SUA	.	.	.	.	.	0
	Rural	3	5	10	15	38	11092
Otago	MUA	4	5	10	15	25	19710
	SUA	3	5	7	10	20	2375
	Rural	3	5	10	15	30	7192
Southland	MUA	4	5	10	10	25	9075
	SUA	2	5	5	10	27	2074
	Rural	3	5	10	20	39	6527
Taranaki	MUA	4	5	10	13	25	9678
	SUA	2	5	5	10	25	3086
	Rural	3	5	10	20	30	6849
Waikato	MUA	4	5	10	15	30	25059
	SUA	3	5	9	15	35	4187
	Rural	3	5	10	17	40	15278
Wellington	MUA	4	5	10	18	35	44359
	SUA	3	5	5	10	25	2441
	Rural	3	5	10	15	35	3501
West Coast	MUA	.	.	.	.	.	0
	SUA	3	5	6	10	21	3719
	Rural	3	5	5	10	30	8440

Note: Only journeys from households with full responses are included.

Table 9: Journey Departure Hour by Region, Urban/Rural in HTS, years 2-12

		Departure Hour																			
Region	Area	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5
Auckland	MUA	94	50	56	29	27	35	38	51	66	106	195	418	634	1238	1931	2727	4328	4457	2626	2428
	SUA	3	0	0	0	0	0	0	3	4	2	16	7	22	55	47	80	93	96	56	91
	Rural	1	0	2	1	1	1	0	0	1	6	4	25	19	46	68	69	102	174	123	131
Bay Of Plenty	MUA	13	19	11	16	14	5	15	8	11	45	28	92	101	241	305	612	1114	1134	781	799
	SUA	1	5	0	1	0	0	0	0	2	0	1	13	16	20	58	95	139	81	81	63
	Rural	2	2	0	0	0	1	0	3	3	7	5	23	20	22	64	200	228	197	165	197
Canterbury	MUA	86	62	73	38	39	33	33	45	45	61	90	207	347	628	997	1746	2557	2855	1932	1935
	SUA	1	9	8	1	1	2	4	3	3	7	12	54	51	102	78	157	265	300	186	244
	Rural	36	23	30	20	12	6	7	9	30	44	72	126	184	335	616	784	1242	1680	921	963
Gisborne	MUA	3	4	6	0	0	0	4	4	5	10	11	44	83	115	149	261	534	547	274	355
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	2	4	0	0	1	0	1	1	4	10	10	19	32	38	68	150	207	140	104	101
Hawke's Bay	MUA	16	15	9	6	9	5	5	9	10	20	26	54	72	168	255	396	716	654	474	473
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	2	1	0	1	0	0	0	0	5	19	16	24	18	52	79	81	197	153	129	113
Manawatu-Wanganui	MUA	7	2	6	3	3	5	6	3	4	14	6	41	56	96	140	298	412	513	262	250
	SUA	3	1	4	2	6	3	2	6	3	4	18	33	36	67	122	148	220	308	174	191
	Rural	11	9	2	2	4	0	3	1	11	9	14	25	47	50	80	223	283	318	226	257
Nels-Marlb-Tas	MUA	4	15	9	2	5	6	10	9	4	17	15	50	32	98	142	261	353	591	334	394
	SUA	7	3	3	5	0	0	2	2	4	4	5	11	32	63	75	113	159	276	149	188
	Rural	2	3	1	2	6	4	3	0	3	5	6	15	27	39	97	188	263	312	224	213
Northland	MUA	1	1	2	1	8	6	2	2	1	13	9	16	46	85	129	226	429	375	227	255
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	4	2	5	5	1	2	3	8	9	10	40	43	72	128	194	328	674	516	365	396
Otago	MUA	14	14	16	7	9	14	13	11	8	12	21	44	86	140	252	482	754	956	572	612
	SUA	2	2	1	0	3	0	0	0	0	1	4	15	20	16	50	51	66	112	78	102
	Rural	12	8	7	5	5	8	5	2	5	6	16	20	23	55	116	197	293	410	214	243
Southland	MUA	8	9	2	11	0	6	11	9	11	13	27	31	56	68	97	241	432	574	240	278
	SUA	0	5	2	1	2	0	3	0	2	1	6	8	3	17	21	45	65	104	64	71
	Rural	6	1	2	0	0	2	2	4	8	14	19	26	42	48	83	198	297	344	180	197
Taranaki	MUA	7	11	3	4	9	3	8	1	14	14	15	28	46	76	138	264	390	499	347	322
	SUA	9	2	2	2	3	0	1	1	6	4	12	21	25	33	34	55	135	163	108	125
	Rural	7	3	6	7	10	4	5	2	7	6	47	18	60	87	115	205	229	335	236	235
Waikato	MUA	29	18	15	7	15	17	16	21	28	32	51	85	144	303	419	821	1307	1167	824	766
	SUA	3	6	2	5	4	0	4	0	2	6	14	20	35	55	75	109	184	285	168	139
	Rural	8	3	13	4	9	0	1	6	20	33	36	78	94	197	328	456	747	850	482	507
Wellington	MUA	55	30	27	21	26	11	13	12	13	19	57	124	251	571	980	1337	1861	2024	1268	1325
	SUA	1	3	2	2	2	2	4	1	4	3	9	7	24	27	17	58	94	112	68	101
	Rural	2	0	1	0	2	0	0	0	2	2	12	23	22	54	56	78	120	148	115	143
West Coast	MUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	SUA	3	0	3	3	4	3	1	0	8	1	10	29	28	62	34	84	103	165	105	105
	Rural	2	4	3	4	6	3	6	3	8	19	22	54	39	88	100	192	243	421	253	236

Note: Only journeys from households with full responses are included. Each column represents a half-hour slot in a 24-hour day. For example, column "0" represents the time between 0:00 and 0:29am.

Table 9: Journey Departure Hour by Region, Urban/Rural in HTS, years 2-12 (continued)

		Departure Hour															
Region	Area	10	10.5	11	11.5	12	12.5	13	13.5	14	14.5	15	15.5	16	16.5	17	17.5
Auckland	MUA	2774	2591	2772	2773	3038	2834	2623	2344	2675	3272	5707	3334	3533	3491	3794	2780
	SUA	63	87	105	85	97	72	69	69	71	86	119	88	92	96	110	78
	Rural	145	119	126	115	87	100	92	66	90	157	167	122	130	132	155	87
Bay Of Plenty	MUA	900	903	918	851	887	825	796	732	816	1057	1428	853	960	918	933	645
	SUA	80	61	70	76	76	92	58	51	54	112	160	79	88	79	86	77
	Rural	158	166	184	133	151	151	123	116	147	178	303	169	155	158	173	118
Canterbury	MUA	2047	2025	2202	2240	2449	2290	2329	2106	2257	2692	3962	2596	2816	2704	2940	2161
	SUA	260	258	249	303	311	291	310	285	308	288	527	304	364	306	362	297
	Rural	974	997	1017	1106	1166	1042	1108	1031	1161	1300	2105	1209	1402	1336	1557	1098
Gisborne	MUA	374	347	352	371	473	386	408	296	335	452	643	400	333	345	528	363
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	111	94	95	124	130	94	141	108	105	152	234	122	113	144	133	94
Hawke's Bay	MUA	490	537	502	560	568	509	558	478	513	616	829	574	640	561	671	474
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	104	109	76	121	87	94	110	115	104	111	235	157	139	160	132	101
Manawatu-Wanganui	MUA	280	338	305	383	389	261	332	281	295	322	573	327	349	378	453	293
	SUA	199	168	209	221	186	198	181	155	169	175	316	232	232	242	238	154
	Rural	239	281	257	235	262	257	268	191	229	248	411	239	285	262	332	183
Nels-Marlb-Tas	MUA	336	368	349	447	443	494	401	384	374	517	717	513	486	515	471	330
	SUA	164	176	196	189	184	202	184	175	157	180	252	167	212	224	221	168
	Rural	174	242	229	226	230	226	244	219	218	269	413	269	281	276	264	187
Northland	MUA	296	219	299	259	309	290	224	245	243	291	495	278	348	298	277	218
	SUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	443	425	458	424	434	433	429	328	403	505	721	405	463	411	492	370
Otago	MUA	653	695	754	781	764	802	767	669	701	836	1285	739	794	926	926	803
	SUA	92	99	94	79	81	70	68	60	83	120	162	129	141	109	99	73
	Rural	260	245	270	243	295	264	303	243	274	249	415	280	276	324	352	300
Southland	MUA	297	309	303	338	423	383	329	288	310	399	656	322	353	350	422	292
	SUA	96	69	89	96	74	80	66	66	74	74	133	87	113	86	83	51
	Rural	216	257	249	238	285	255	243	236	215	252	497	246	290	303	289	210
Taranaki	MUA	324	363	379	346	465	415	333	323	378	394	614	360	424	426	450	302
	SUA	127	117	128	125	120	127	142	80	100	117	189	111	120	139	137	80
	Rural	286	269	253	269	263	255	272	224	265	273	416	294	294	258	301	241
Waikato	MUA	818	830	910	991	972	925	824	734	813	926	1589	998	1187	1043	1208	903
	SUA	135	140	155	159	148	181	160	146	126	178	274	167	165	184	166	110
	Rural	583	526	538	577	560	522	574	456	535	633	1112	592	624	604	687	516
Wellington	MUA	1364	1413	1527	1596	1810	1714	1736	1448	1485	1644	2713	1809	1962	1983	2145	1692
	SUA	98	94	97	107	100	105	115	102	100	80	125	98	82	89	94	96
	Rural	167	176	187	156	142	128	127	113	114	110	205	160	140	150	139	102
West Coast	MUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	SUA	130	133	158	157	150	159	123	111	127	183	253	166	181	155	146	126
	Rural	289	380	383	356	388	404	336	305	317	354	639	374	347	346	350	228

Note: Only journeys from households with full responses are included.

Table 9: Journey Departure Hour by Region, Urban/Rural in HTS, years 2-12 (continued)

		Departure Hour											
Region	Area	18	18.5	19	19.5	20	20.5	21	21.5	22	22.5	23	23.5
Auckland	MUA	1197	896	884	669	520	399	316	262				
	SUA	43	56	46	46	26	14	12	22	7	3	2	3
	Rural	68	44	40	21	34	23	29	15	16	9	5	4
Bay Of Plenty	MUA	503	364	340	257	236	186	131	110	95	86	56	49
	SUA	44	60	14	32	15	26	42	10	5	8	3	4
	Rural	116	74	67	45	53	24	24	15	25	15	2	3
Canterbury	MUA	1976	1365	1295	968	906	724	596	481	355	297	273	203
	SUA	200	153	154	111	84	91	67	48	31	25	30	17
	Rural	849	680	545	386	343	254	238	220	151	123	74	61
Gisborne	MUA	252	142	155	120	81	68	79	43	60	40	9	18
	SUA	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	54	47	49	31	16	16	12	8	6	7	11	5
Hawke's Bay	MUA	332	286	237	162	148	90	97	128	63	47	32	29
	SUA	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	92	47	42	28	25	17	23	15	2	7	6	9
Manawatu-Wanganui	MUA	252	175	159	110	100	74	79	45	41	43	32	6
	SUA	120	101	82	51	46	44	38	32	24	16	12	5
	Rural	171	113	81	58	85	55	48	20	28	28	13	18
Nels-Marlb-Tas	MUA	236	221	206	111	125	123	90	74	42	24	30	18
	SUA	141	76	80	76	58	50	35	20	18	24	16	5
	Rural	137	119	83	97	68	43	27	38	34	9	10	8
Northland	MUA	198	116	111	117	72	72	29	48	27	26	12	8
	SUA	0	0	0	0	0	0	0	0	0	0	0	0
	Rural	285	190	144	141	88	82	60	55	28	34	16	25
Otago	MUA	555	424	401	269	280	185	171	168	149	74	78	40
	SUA	69	46	41	38	18	23	12	10	9	13	11	3
	Rural	197	131	127	129	80	74	53	35	51	33	24	15
Southland	MUA	228	160	185	143	133	72	64	55	35	50	31	21
	SUA	59	58	28	38	38	31	21	5	17	4	13	5
	Rural	173	121	128	82	67	62	43	36	19	8	21	14
Taranaki	MUA	249	186	164	133	105	75	104	51	28	37	24	28
	SUA	87	74	48	32	34	17	22	22	26	14	4	6
	Rural	203	136	118	65	60	48	41	37	31	24	16	16
Waikato	MUA	718	551	491	377	277	230	197	161	128	95	58	59
	SUA	103	101	53	54	40	39	17	13	28	27	2	0
	Rural	410	329	213	182	171	116	117	90	55	27	39	21
Wellington	MUA	1294	1045	869	622	602	432	387	352	240	197	139	114
	SUA	69	56	37	40	32	19	19	13	15	10	2	6
	Rural	88	64	78	33	36	22	17	35	11	14	7	1
West Coast	MUA	0	0	0	0	0	0	0	0	0	0	0	0
	SUA	121	75	88	61	36	29	22	39	18	12	6	3
	Rural	181	156	159	71	76	77	89	51	46	24	18	13

Note: Only journeys from households with full responses are included.

Table 10 investigates whether multi-leg journeys are common in the HTS. This will guide us in our choice of dependent variable in any DCA, for example whether it is necessary to include multi-mode choices. This table shows that the vast majority of journeys in the HTS are single-leg journeys. There are very few journeys with more than two trips. Table 11 further explores the travel mode combination among journeys with two trips. The most common ones are "walk then drive" and "drive then walk". One should note that it is possible that many very short (less than 5 minutes) legs are under-reported, for example, a short drive to the train station park-and-ride or a short walk to the destination from the car park.

The following two tables, Tables 12 and 13, explore two characteristics of vehicles owned by households surveyed, vehicle age and type. These vehicle characteristics are potentially useful for constructing a measure of ride quality, which is one of the factors that determine travel mode choice. The distribution of vehicle age (in terms of percentiles) is extremely similar across all regions. There is also not much noticeable difference between urban and rural areas. In terms of vehicle type, the HTS allows for seven categories; not surprisingly, the most reported type is "car/ station wagon".

Table 14 explores the data on parking in the HTS. This is potentially useful for constructing parking costs to be incorporated in the DCA. When the travel mode of a trip is driving a vehicle, the HTS includes questions on whether the vehicle is parked, and if yes, which one of the six types of parking is used.

One important drawback of the HTS "old survey" is that no route information is collected; instead, we only know the origin and destination of each trip. (The "new survey" that uses GPS units is likely to remedy this.) Without information on route, we cannot reliably incorporate external data on traffic volume or congestion as characteristics of the chosen mode choice. For example, if we observe that a journey between two suburbs taken by car has a significantly longer duration than the free-flow travel time, we cannot deduce whether it is due to the driver choosing to take slower local roads, or that the driver has indeed taken the state highway, but it is was congested. Having this information will allow us to have much better estimation on preference for travel time and its reliability, both of which are heavily dependent on route information.

Table 15 is a preliminary investigation on whether we can reliably deduce the route taken, for at least a small portion of journeys taken. Our assumption here is that, if both

Table 10: Number of Journeys by Number of Legs, Region, Urban/Rural in HTS, years 2-12

Region	Area	Number of legs 1	2	3	4	5	6	7	8
Auckland	MUA	81397	1794	1258	178	51	9	2	4
	SUA	2204	24	13	1	0	0	0	0
	Rural	2897	57	15	1	1	0	1	0
Bay Of Plenty	MUA	21617	486	88	6	2	0	0	0
	SUA	2080	58	0	0	0	0	0	0
	Rural	4026	146	13	0	0	0	0	0
Canterbury	MUA	59924	1914	1015	179	22	8	1	1
	SUA	7158	334	25	5	0	0	0	0
	Rural	29479	925	219	44	5	0	0	1
Gisborne	MUA	9366	495	18	1	2	0	0	0
	SUA	0	0	0	0	0	0	0	0
	Rural	3025	115	8	0	0	0	0	0
Hawke's Bay	MUA	13693	378	49	3	0	0	0	0
	SUA	0	0	0	0	0	0	0	0
	Rural	3068	89	1	0	0	0	0	0
Manawatu-Wanganui	MUA	8567	199	30	6	0	0	0	0
	SUA	5104	79	13	1	0	0	0	0
	Rural	6313	146	13	0	0	0	0	0
Nels-Marlb-Tas	MUA	10388	377	30	1	0	0	0	0
	SUA	4554	187	8	1	0	1	0	0
	Rural	5764	261	26	2	0	0	0	0
Northland	MUA	6664	532	59	4	0	0	0	0
	SUA	0	0	0	0	0	0	0	0
	Rural	10525	536	29	7	0	0	0	0
Otago	MUA	18588	964	163	7	3	1	0	0
	SUA	2318	55	2	0	0	0	0	0
	Rural	6974	195	22	1	0	0	0	0
Southland	MUA	8916	143	13	3	0	0	0	0
	SUA	2067	7	0	0	0	0	0	0
	Rural	6307	211	9	1	0	0	0	0
Taranaki	MUA	9439	215	20	2	2	1	0	0
	SUA	3019	57	6	4	0	0	0	0
	Rural	6677	161	12	2	0	0	0	0
Waikato	MUA	24273	664	143	16	2	0	0	0
	SUA	4070	102	12	2	1	0	0	0
	Rural	14815	438	22	5	1	0	0	0
Wellington	MUA	40665	2133	1286	230	34	7	4	0
	SUA	2371	48	21	1	0	0	0	0
	Rural	3322	126	37	12	4	1	0	0
West Coast	MUA	0	0	0	0	0	0	0	0
	SUA	3605	107	7	0	0	0	0	0
	Rural	8252	205	6	0	0	0	0	0

Note: Only trips from households with full responses are included.

Table 11: Number of 2-Leg Journeys by Travel Mode Combinations, Region, Urban/Rural in HTS, years 2-12

		Travel mode combination									
Region	Area	Walk then VD	VD then Walk	Walk then VP	Walk then Bus	VP then Walk	Bus then Walk	VD then VD	VP then Bus	VP then VP	Bus then VP
Auckland	MUA	306	271	170	271	122	213	95	43	39	28
	SUA	2	2	1	0	0	1	6	3	1	3
	Rural	11	8	6	7	5	8	1	2	1	3
Bay Of Plenty	MUA	110	79	62	68	37	54	5	11	20	8
	SUA	4	2	1	15	6	16	2	1	8	1
	Rural	20	11	13	31	7	25	2	15	4	17
Canterbury	MUA	453	409	182	255	146	155	55	13	35	26
	SUA	110	104	38	22	30	14	1	1	1	2
	Rural	226	196	116	86	77	74	14	16	15	13
Gisborne	MUA	135	158	65	12	75	9	9	1	11	2
	SUA	0	0	0	0	0	0	0	0	0	0
	Rural	18	24	12	18	5	12	1	4	3	1
Hawke's Bay	MUA	77	59	55	53	43	47	5	11	0	9
	SUA	0	0	0	0	0	0	0	0	0	0
	Rural	18	20	11	4	10	10	1	6	0	5
Manawatu-Wanganui	MUA	54	40	18	26	11	23	3	3	2	1
	SUA	12	6	14	2	8	3	3	4	8	2
	Rural	31	18	17	12	13	10	9	11	5	11
Nels-Marlb-Tas	MUA	128	97	44	24	29	15	11	0	3	1
	SUA	59	57	30	3	20	5	2	1	0	0
	Rural	76	44	27	17	13	15	6	17	1	6
Northland	MUA	144	168	74	12	75	17	12	2	5	5
	SUA	0	0	0	0	0	0	0	0	0	0
	Rural	117	110	73	50	49	48	8	27	16	18
Otago	MUA	320	281	112	56	79	30	10	7	8	4
	SUA	14	12	9	5	6	3	0	0	1	0
	Rural	58	55	14	19	9	8	3	7	0	0
Southland	MUA	27	20	24	15	2	14	4	4	7	1
	SUA	4	0	1	0	0	0	0	2	0	0
	Rural	26	18	16	27	14	26	9	19	2	20
Taranaki	MUA	68	55	34	6	16	4	10	1	3	2
	SUA	17	7	10	2	10	2	4	0	2	2
	Rural	30	25	23	16	6	10	9	5	6	6
Waikato	MUA	190	164	71	41	54	42	8	18	12	13
	SUA	31	19	18	6	3	9	3	4	0	1
	Rural	105	81	50	26	39	19	10	27	16	26
Wellington	MUA	538	495	264	152	183	121	132	41	30	21
	SUA	17	10	3	0	0	0	15	0	1	0
	Rural	34	31	10	3	7	4	15	6	7	4
West Coast	MUA	0	0	0	0	0	0	0	0	0	0
	SUA	33	24	18	0	16	1	2	1	1	0
	Rural	40	24	35	14	16	8	16	10	13	7

Note: Only journeys consisting of 2 trips, from households with full responses are included. In travel modes, "VD" stands for "vehicle driver"; "VP" stands for "vehicle passenger". The ten columns shown are the ten most common travel mode combinations in the HTS.

Table 12: Percentiles of Vehicle Age by Region, Urban/Rural in HTS, years 2-12

		Percentiles				
Region	Area	10th	30th	50th	70th	90th
Auckland	MUA	3	8	11	14	18
	SUA	3	7	12	15	19
	Rural	3	8	11	15	20
Bay Of Plenty	MUA	3	7	11	15	19
	SUA	4	8	12	14	18
	Rural	4	8	11	15	21
Canterbury	MUA	4	9	13	16	21
	SUA	3	8	13	16	22
	Rural	3	8	12.50	16	22
Gisborne	MUA	3	8	12	15	21
	SUA	.	.	.	.	.
	Rural	3	7	11	15	22
Hawke's Bay	MUA	3	8	12	15	21
	SUA	.	.	.	.	.
	Rural	2	7	11	15	21
Manawatu-Wanganui	MUA	2	7	10	14	20
	SUA	3	9	13	16	20
	Rural	3	7	11	15	20
Nels-Marlb-Tas	MUA	4	9	13	17	23
	SUA	3	8	12	15	22
	Rural	3	9	13	17	23
Northland	MUA	2	8	11	15	20
	SUA	.	.	.	.	.
	Rural	3	8	12	15	21
Otago	MUA	4	9	13	16	22
	SUA	3	9	12	16	24
	Rural	2	7	11	15	21
Southland	MUA	4	9	12	16	22
	SUA	3	9	12	16	23
	Rural	2	6	10	14	21
Taranaki	MUA	2	7	11	14	19
	SUA	3	7	11	15	21
	Rural	3	8	11	15	21
Waikato	MUA	3	7	11	14	20
	SUA	3	7	12	15	20
	Rural	2	7	11	15	21
Wellington	MUA	3	7	10	14	18
	SUA	3	8	12	16	22
	Rural	3	8	11	15	22
West Coast	MUA	.	.	.	.	.
	SUA	2	6	11	14	19
	Rural	3	9	13	16	21

Note: Only journeys from households with full responses are included.

Table 13: Number of Vehicles by Type, Region, Urban/Rural in HTS, years 2-12

		Vehicle Type						
Region	Area	1	2	3	4	5	6	7
	MUA	6974	740	990	105	3	151	44
Auckland	SUA	178	18	31	5	0	1	4
	Rural	290	60	58	22	0	4	11
	MUA	1969	310	256	36	2	56	25
Bay Of Plenty	SUA	130	45	32	6	0	8	2
	Rural	337	89	96	20	0	11	6
	MUA	5328	493	807	61	5	141	52
Canterbury	SUA	613	80	155	25	0	16	13
	Rural	2500	354	619	70	0	115	67
	MUA	635	154	130	31	0	39	5
Gisborne	SUA	.	.	.	.	.	.	.
	Rural	237	92	76	28	1	19	23
	MUA	1082	273	165	22	0	52	19
Hawke's Bay	SUA	.	.	.	.	.	.	.
	Rural	217	76	80	13	0	50	22
	MUA	740	124	81	21	0	21	10
Manawatu-Wanganui	SUA	447	80	70	17	0	33	5
	Rural	499	149	124	34	1	32	14
	MUA	747	163	147	24	2	60	18
Nels-Marlb-Tas	SUA	357	78	89	14	0	24	14
	Rural	453	155	164	27	0	25	31
	MUA	577	99	87	22	0	17	34
Northland	SUA	.	.	.	.	.	.	.
	Rural	867	265	293	70	1	51	145
	MUA	1571	216	214	25	1	42	13
Otago	SUA	185	38	33	3	0	7	4
	Rural	514	110	208	23	1	26	18
	MUA	741	99	151	11	2	31	5
Southland	SUA	196	27	34	5	0	8	0
	Rural	571	126	232	63	2	34	33
	MUA	732	118	94	15	0	28	15
Taranaki	SUA	237	75	19	5	0	20	6
	Rural	616	179	111	30	1	61	30
	MUA	2166	365	270	60	0	71	32
Waikato	SUA	311	55	47	19	1	6	8
	Rural	1287	342	292	95	0	81	68
	MUA	3412	420	419	49	0	140	29
Wellington	SUA	197	33	11	11	0	3	5
	Rural	303	71	31	15	0	8	15
	MUA	.	.	.	.	.	.	.
West Coast	SUA	259	65	71	8	0	14	8
	Rural	567	160	250	21	1	37	43

Note: Only journeys from households with full responses are included. Vehicle type codes: 1 = car / station wagon; 2 = van / ute / passenger van; 3 = four wheel drive; 4 = truck; 5 = taxi; 6 = motor bike; 7 = other.

Table 14: Parking Type by Region, Urban/Rural in HTS, years 2-12

Region	Area	Not parked	Residential	Private	Off street Public	Time limit	No time limit	Other
	MUA	2532	20140	12022	6851	1084	4234	39
Auckland	SUA	32	568	422	168	46	38	4
	Rural	75	782	493	194	31	143	2
	MUA	482	5427	3542	1490	442	1137	10
Bay Of Plenty	SUA	40	465	301	105	68	87	0
	Rural	55	1133	748	138	64	281	14
	MUA	1504	13203	10184	2935	798	5053	11
Canterbury	SUA	301	1834	1407	180	168	784	2
	Rural	783	6703	4971	1369	279	2369	16
	MUA	185	2231	1225	391	217	944	1
Gisborne	SUA	0	0	0	0	0	0	0
	Rural	64	803	492	130	75	241	1
	MUA	439	3243	2446	505	393	879	6
Hawke's Bay	SUA	0	0	0	0	0	0	0
	Rural	102	773	583	107	97	179	3
	MUA	285	2094	1573	376	319	541	20
Manawatu-Wanganui	SUA	168	1391	942	122	112	408	5
	Rural	180	1734	915	261	117	540	26
	MUA	360	2482	1701	564	170	641	12
Nels-Marlb-Tas	SUA	113	1214	745	336	95	226	25
	Rural	209	1669	979	303	120	276	12
	MUA	192	1653	1193	381	103	424	0
Northland	SUA	0	0	0	0	0	0	0
	Rural	242	2944	1945	575	115	470	10
	MUA	422	3403	2930	411	758	2484	14
Otago	SUA	28	501	445	44	88	256	2
	Rural	120	1694	1216	264	193	659	8
	MUA	293	2192	947	748	284	739	8
Southland	SUA	39	516	242	139	64	186	1
	Rural	131	1926	748	496	159	531	21
	MUA	206	2161	1768	393	263	723	1
Taranaki	SUA	74	838	449	107	60	223	5
	Rural	143	1940	1053	228	122	444	2
	MUA	616	6286	4801	1242	377	1224	30
Waikato	SUA	101	980	576	252	111	221	9
	Rural	345	4146	2487	832	225	990	24
	MUA	1174	8351	4813	3301	1165	3667	84
Wellington	SUA	80	624	292	271	43	196	13
	Rural	109	916	451	248	62	302	3
	MUA	0	0	0	0	0	0	0
West Coast	SUA	103	918	505	232	177	210	14
	Rural	151	2057	1038	330	293	608	53

Note: Only trips from households with full responses, and trips with travel mode equals 1 (vehicle driver), are included.

the origin and destination of a journey mention the same street name, it can reasonably be safely deduced that the actual route taken is, in fact, on that street. (Of course, this is far from the only possible assumption that can be made to generate a possible route.) Unfortunately, Table 15 shows that likely very few journeys satisfy this criterion. The most commonly mentioned street name among all origins and destinations, Colombo Street in Canterbury, was mentioned only 1202 times in the HTS. The number of journeys with both origin and destination on Colombo Street is likely to be a magnitude lower than that. Thus, it is unlikely that the existing HTS will provide usable route choice information.

Table 15: Most Commonly Mentioned Street Names in HTS, years 2-12

Street Name	Region	N
Colombo St	Canterbury	1202
Riccarton Road	Canterbury	1120
Great South Road	Auckland	985
Gladstone Road	Gisborne	812
High St	Canterbury	742
Main North road	Canterbury	738
Gerald St	Canterbury	532
Queen St	Auckland	501
Moorhouse Avenue	Canterbury	500
Cameron Road	Bay of Plenty	468
Victoria St	Waikato	437
Ferry Road	Canterbury	427
Great North Road	Auckland	417
Blenheim Road	Canterbury	413
SH 1S Tay St	Southland	400
George St	Otago	383
Rolleston Drive	Canterbury	378
Cashel St	Canterbury	375
Victoria St	Canterbury	375
Te Rapa Road	Waikato	362

Note: Only addresses in trips from households with full responses are included.

4 Implementing DCA with HTS and Other Data

4.1 Choice Variable

The HTS contains rich enough data to model transport users' choice of transport mode (i.e. car, bus, cycle, etc), which is a common type of choice modelled in transport studies. It does not contain enough data to model route choices, or activity choices, and it is unlikely that other data sources could be used to model these other types of choices. Future HTS data – using recently-incorporated GPS-based journey measures – will better enable consideration of other transport choices, such as route choice.

4.2 Trip Types

The HTS contains sufficient data to analyse mode choices for urban trips of short duration. It is not rich enough to model inter-city mode choice decisions, and hence cannot be used to model aviation versus land-based passenger vehicles. It is possibly rich enough to include commuter trains and ferries in any mode choice analysis, if only for Wellington and Auckland.

The HTS is also rich enough for individual trips to be chain-linked into journeys. This enables estimation of a DCA that explains mode choice based on respondent-level journey data. Where a journey comprises trips involving more than one mode, the mode for that journey could be taken to be the "dominant mode", for example based on total distance travelled for each mode on that journey.

4.3 Non-Chosen Alternatives

An important limitation of the HTS is that it contains data only on transport alternatives actually chosen by survey respondents. It does not, therefore, contain information on transport alternatives available to but not chosen by each respondent. When using respondent-level data, it is necessary that the researcher has data on both chosen and at least one non-chosen alternative.[47] However, the HTS contains origin and destination data, trip lengths, and car and cycle availability, for each respondent. Hence, it should be possible to reverse-engineer at least one non-chosen alternative per survey response

[47] Techniques exist for estimating respondents' preferences using aggregated data, in which case non-chosen alternatives are not required (e.g. see Small and Verhoef (2007) for transport-related references, or Girotti and Meade (2015) for a more general summary of the methodology). However, in this case it remains necessary to identify an "outside option" – e.g. what travellers choose if they decide not to travel at all.

– for many respondents at least. Possible approaches include using public transport databases for alternatives to private vehicle use, or private vehicle use (including cycles or walking for shorter trips) as alternatives to public transport use.

4.4 Travel Cost

Another important limitation of the HTS is that it does not contain any comprehensive measure of travel cost. Such a financial variable is necessary in order for WTPs to be estimated. It should be possible, however, to supplement the HTS using "engineering-based" estimates of travel costs.[48] Combined also with estimates of parking costs, this should provide meaningful cost information enabling the estimation of WTP for non-monetary transport attributes. Vehicle operating cost (VOC) estimates in NZTA's Economic Evaluation Manual (EEM) would be a suitable source for such estimates by vehicle class for travel by private vehicle.[49] Travel cost is often deflated by income, based on theoretical models of transport user behaviour.[50]

For public transport, fare schedules and zones could be combined with HTS origin and destination data to estimate respondent's public transport costs. For active modes such as cycling, travel costs could be estimated to be nil, or based on any publicly available direct cost estimates.

4.5 WTP for Travel Time Savings – VOT

The HTS contains respondent-level trip durations. When trips are chain-linked into journeys, this also provides information on different elements of travel time, such as transit time between modes. However, the HTS does not provide detailed breakdowns of travel time within a mode, such as free-flow or congested travel time when travelling by car.[51] Hence, it will not enable estimation of separate VOTs for each such sub-component of travel time.

[48] Small and Verhoef (2007) note that it is common practice to use such estimates to improve transport datasets. Cost estimates could be interacted with indices of fuel costs to introduce time-series variation in engineering costs, in order to reflect the impact of changing fuel prices. Also, travel cost is often deflated by income, based on theoretical models of transport user behaviour – see, for example, Small (2012), Hensher (2011), or Chapter 2 of Small and Verhoef (2007).

[49] NZTA (2016). In particular, see Appendix A5.

[50] See, for example, Small (2012), Hensher (2011), or Chapter 2 of Small and Verhoef (2007).

[51] Small and Verhoef (2007), and Hensher (2011), discuss evidence that VOT varies depending on whether travel is free-flowing or congested.

As well as trip and journey durations, the HTS contains other variables such as travel mode, origin and destination, and mode and respondent characteristics, relevant to VOT estimation. Combined with engineering-based estimates of travel cost, this data should enable estimation of transport users' WTP for travel time savings (i.e. VOT).[52]

4.6 WTP for Travel Time Reliability/Variability – VOR

Standard approaches for estimating the Value of Reliability (VOR) in transport studies measure reliability in terms of the spread of travel times (e.g. travel time standard deviation, or differences between various travel time deciles).[53] Reflecting the asymmetric impact of travel time variability on transport users (i.e. on average, arriving late is more undesirable than arriving early), a common measure of reliability is the 90th (or 80th) percentile of travel times less the median travel time.[54] Such asymmetric measures of reliability have been found to be superior to symmetric measures such as travel time standard deviation. This is because transport-users include greater buffer times when choosing departure times, to reduce the risk of arriving late at their destination due to travel time variability.

Such measures require access to a relevant distribution of travel times, ideally for a given route. It should be possible to estimate such measures of spread using the HTS, given multiple observations for different modes and trip lengths, controlling for trip length.[55] Combined with engineering-based estimates of travel cost, this data should enable estimation of transport users' WTP for travel time reliability (i.e. VOR).

4.7 WTP for Safety – VOS

The HTS contains no safety data for each mode. While trip length or travel time might be used to proxy for trip safety, that would confound estimation of

[52] In this research, focus is on behavioural models of travel time. An alternative approach is to use productivity models, such as those summarised in Hensher (2011). For VOT, the behavioural approach is most commonly applied, and is appropriate for analysing the private or work-related transport choices of individual transport users. Productivity-based models can be applied for business-related trips.

[53] Bhouri et al. (2012) summarise common reliability measures.

[54] Small (2012), Small and Verhoef (2007).

[55] Trip length must be controlled for since longer trips will, on average, involve greater travel times. Hence the travel time distribution will vary based on trip length. It is not necessary to control also for modes, despite different modes having different travel times, since this is one of the relevant determinants of transport users' mode choice.

transport users' preference for trip length, or for VOT itself. Hence estimating WTP for safety – or Value of Safety (VOS) – requires other approaches. Usefully, detailed historical data on both fatal and non-fatal accident rates, by mode, are publicly available.[56] Such accident rates are either per unit of time travelled, or per unit of distance travelled. They can each be scaled up to total accident risk per journey using the journey time and distance data already in the HTS.

Recent DCA studies of VOS measure safety in terms of risk of fatal accidents.[57] They emphasise the trade-off between lower speeds (i.e. greater travel times) and higher accident rates. Traffic density is a variable of interest in such studies, defined as the average number of vehicles on a road per unit of distance at a given time. In practice, traffic density is approximated using finely-specified sub-components of travel time. However, as discussed above, those sub-components are not available in the HTS. Hence, NZTA data on traffic density such as from the *National Telemetry Series* might instead be used.

More generally, transport mode safety might be more broadly defined. In particular, published data on:

- Non-fatal accidents could also be incorporated alongside fatal accidents as an explanatory variable; and
- Crime data such as numbers of reported criminal incidents could be correlated with journey origin and destination data to measure risk of crime against travellers (e.g. of theft or assault at bus stops, train stations or parking buildings).[58]

Since estimating VOS is relatively new in transport DCA studies, it is helpful that there is a variety of data with which to trial alternative approaches.

4.8 WTP for Ride Quality – VOQ

The HTS contains only limited data of relevance for estimating transport users' WTP for ride quality – what can be called Value of Quality (VOQ). In particular, while it contains information such as vehicle type and age for

[56]For example, see Ministry of Transport (2015a, 2015b, 2015c).

[57]See Steimetz (2008), Hensher (2006).

[58]New Zealand Police (2016) describes available historical reported crime data. McGirr et al. (2013) describe how to create regional crime statistics from historical reported crime data. Statistics New Zealand now publishes monthly crime statistics, including at the community level for each police boundary. Such community-level data produces crime rates for health, education, justice, religious, and transport facilities and services as well as roads, streets, carparks, footpaths, and open space.

car trips – each of which can proxy for quality – it contains no comparable information for other modes such as buses or cycles. Moreover, "ride quality" is hard to compare across transport modes – e.g. driver friendliness and punctuality are relevant quality attributes for public transport, but not for cycling or private vehicle trips.

It should be possible to construct proxies for ride quality, but only for selected modes – in particular, motorised vehicles. For example, (average) fleet age could be obtained for public transport fleets, as a proxy for vehicle quality. This would provide a quality measure that could then be compared with vehicle age for private vehicles. Indicators could also be included for whether or not the vehicle is private, shared or public, and also for the particular type of passive mode. These would then provide quality indicators relative to passive modes (assuming they are attributed no quality indicator). For modes such as walking, cycles, scooters and motorcycles, a dummy could be included for weather-exposure. [59]

It is likely that experimentation will be required to identify meaningful and significant ride quality indicators that can be compared across modes. Even if such quality indicators cannot be identified and adequately measured, it is possible to specify DCA models so as to measure average preferences for each mode.[60] These average preferences measure transport user preferences for non-measured attributes, which include non-measured quality attributes – relative to the average preference for a selected alternative.[61] Hence, it should be possible to measure VOQ in a relative and/or partial sense, if not in an absolute sense, using HTS data supplemented with other data sources.

4.9 HTS Sample Weights

Finally, as noted in Section 3.2, the HTS includes sampling weights, as well as non-response weights and post-stratification weights. These are handled in the usual way for DCA, by re-weighting choice probabilities in the log-likelihood function used for estimation.[62]

[59] Another possibility is to locate existing research on user evaluations of the quality of different transport modes.

[60] Formally, this involves the inclusion of alternative-specific constants for all alternatives except one in the DCA model – e.g. see Train (2009), or Small and Verhoef (2007).

[61] Specifically, including alternative-specific constants measures differences in transport users' average preference for any given mode relative to their average preference for the mode which is not assigned an alternative-specific constant. For example, if all modes except cycling are assigned alternative-specific constants, then estimates of those constants measure the average preference for that mode relative to cycling.

[62] References are provided in Small and Verhoef (2007). Steimetz (2008) provides details in an appendix.

5 Using Discrete Choice Outputs for GPS Activity Class Analyses

As noted in earlier sections, VOT estimated using DCA has found application in transport analyses since its first use in forecasting demand for a light-rail system in San Francisco Bay.[63] Having a reliable estimate of VOT – in monetary terms – enables economic analysis of roading or public transport improvements. For example, these include cost-benefit analyses of improved public transport frequency or lane improvements on roads, or analysis of the business case for toll roads.

Increasingly, VOR is also being incorporated into transport analyses.[64] For example, a network enhancement could be evaluated using both VOT and VOR as follows:[65]

$$\begin{array}{c}\text{Value of time}\\ \text{saving from net-}\\ \text{work enhancement}\end{array} = \begin{array}{c}\text{Journey time}\\ \text{saving}\end{array} \times VOT + \begin{array}{c}\text{Buffer time}\\ \text{saving}\end{array} \times VOR$$

In particular, the Economic Evaluation Manual published by NZTA provides a range of standardised economic analysis frameworks for evaluating transport-related projects.[66] Figure 1 provides an overview of how VOT and VOR are used by NZTA for such evaluations.[67] Here, "Travel time cost savings" refers to VOT, while "Journey time reliability benefits" refers to VOR. From this table it can also be seen that VOS is also relevant, such as for "Crash cost savings".

Section 4 of NZTA (2016) provides details of how such measures are applied in evaluations of:

- Roading activities;
- Transport demand management;
- Transport services;
- Walking and cycling; and

[63] Bay Area Rapid Transit, or BART, as described in McFadden (1974). McFadden (2001) reported that official forecasts at the time were a 15% demand share for the new mode. The DCA approach predicted 6.3%, while the actual mode share was 6.2%.

[64] Section 3 of OECD/ITF (2010) surveys how VOR has been incorporated in transport-related cost-benefit analyses in various countries, including New Zealand.

[65] OECD/ITF (2010), p. 68. van Oort (2016) shows how to include VOR in cost-benefit analyses of public transport improvements.

[66] NZTA (2016).

[67] NZTA (2016), p. 2–8.

Figure 1: NZTA Transport Applications of VOT and VOR

Benefit type	Road	Transport demand management	Transport services	Walking and cycling	Education promotion and marketing	Parking and land use	Private sector financing and road tolling
Travel time cost savings	✓	✓	✓	✓	✓		
Vehicle operating cost savings	✓	✓	✓	✓	✓		
Crash cost savings	✓	✓		✓	✓	✓	
Seal extension benefits	✓						
Driver frustration reduction benefits	✓						
Risk reduction benefits	✓	✓	✓		✓		✓
Vehicle emission reduction benefits	✓					✓	
Other external benefits	✓	✓	✓	✓	✓	✓	✓
Mode change benefits		✓	✓	✓	✓		
Walking and cycling health benefits		✓		✓			
Walking and cycling cost savings		✓		✓	✓		
Transport service user benefits			✓			✓	
Parking user cost savings		✓			✓	✓	
Journey time reliability benefits	✓	✓	✓	✓			
Wider economic benefits	✓		✓				
National strategic factors	✓	✓	✓		✓		

- Private sector financing and road tolling.

The Ministry's GPS output classes are repeated from Section 1 here:

- State highway investment and maintenance;
- Local roads investment and maintenance;
- Regional improvements and Auckland;
- Public transport;
- Walking and cycling;
- Road policing and safety promotion; and
- Investment management.

There are clear overlaps between these output classes and many of the NZTA applications. Hence NZTA's methodologies for applying measures such as VOT, VOR and VOS illustrate how they might be applied in analyses relating to the Ministry's GPS output classes.

6 Recommendations and Conclusions

We recommend that:

- The HTS be supplemented with data from other sources such as those indicated above, to enable DCA to be used to estimate transport users' trade-offs for the transport attributes of interest;
- A DCA model of dominant mode choice for journeys (i.e. chain-linked trips) be estimated, including measures of direct travel cost, travel time, travel time dispersion, safety and quality as explanatory variables to enable estimation of VOT, VOR, VOS and VOQ;[68]
- The model be restricted to urban rather than long-distance travel, and to personal (including work) rather than business-related trips;
- A mixed logit model be estimated, using traveller-level data (including demographics), to estimate how WTPs for transport attributes vary by type of transport user;
- The model be estimated using the full traveller-level dataset, with explanatory variables included for journey location (e.g. urban, rural, etc), and separate models estimated if those explanatory variables indicate significant differences by location; and
- Various measures of travel safety and ride quality be tried, since their measurement is less settled in the literature, and perhaps only for motorised vehicles.

More specifically, we recommend that a mixed logit DCA be implemented with an indirect utility specification along the following lines (defining V_{ij} as the utility derived by traveller i from choosing (dominant) mode j):

$$V_{ij} = \alpha_{ij} + \beta_{ijc}\left(\frac{c_j}{w_i}\right) + \beta_{ijT}T_j + \beta_{ijR}R_j + \beta_{ijS}S_j$$
$$+\beta_{ijTS}\left(T_j \times S_j\right) + \beta_{ijQ}Q_j + \beta_{ij}X_j + \epsilon_{ij}$$

Here:

- α_{ij} is a mode-specific constant (to be estimated) that is allowed to vary by traveller (i.e. we allow for different types of traveller to have different average levels of preference for each mode);

[68] Other relevant explanatory variables include, for example, journey length, purpose, timing and location.

- c_j is the direct cost of mode j, normalised by traveller j's income (or log income) w_i;
- T_j is the all-up travel time for mode j;
- R_j is the travel time variability of mode j, measured as the difference between the 90th percentile and median travel time for journeys of comparable length on that mode;
- S_j is the safety of mode j (perhaps using multiple safety measures);
- $T_j \times S_j$ is the interaction of travel time and safety of mode j;
- Q_j is the ride quality of mode j (perhaps using multiple quality measures);
- X_j are other possible attributes of mode j (e.g. journey length, urban/rural location, public transport dummy, etc);
- β_{ijc}, β_{ijT}, β_{ijR}, β_{ijQ}, β_{ijS}, β_{iTS} and β_{ij} are traveller- and (possibly) mode-specific taste parameters to be estimated; and
- $\epsilon_{ij}$is an iid Type 1 Extreme Value error term that captures consumer heterogeneity not explained by the traveller-specific taste parameters.

The time-safety interaction term proposed above is intended to provide a reduced-form expression for the possible trade-off between travel time and travel safety suggested by theory. Note that all parameters have been specified as traveller-specific in the first instance, to allow for taste-variation across consumers for each travel attribute. In practice some attributes are deliberately chosen to be non-random.[69]

If estimation of a mode choice model using disaggregated data should prove problematic, it is also possible to try estimating the desired WTPs using aggregated data and estimation techniques appropriate to such data. Alternatively, a narrower version of the model could be estimated using disaggregated data, such as a route choice model where the observed data includes known bottlenecks (e.g. Auckland Harbour Bridge, or Hutt Motorway).[70]

[69]For example, the travel cost coefficient β_{ijc} is often simplified to β_{jc}, which is constant across travellers. This is to ensure that WTP estimates are not poorly defined should estimates of this coefficient be close to zero or negative, given it is the common denominator in all required WTP estimates. Otherwise, researchers typically select a subset of taste coefficients to treat as random, based on theory, but also for practical considerations (e.g. parsimony, and to enable statistically significant estimates to be produced by a dataset of given quality).

[70]Such bottlenecks naturally define routes, even if the existing HTS data does not include explicit route data.

We recommend attempting to supplement the HTS data, despite it not being already sufficient for DCA, because it represents an important record of actual travel decisions in New Zealand. This means it is currently the only comprehensive New Zealand dataset containing information on actual transport user choices that might be used to estimate their preferences for different transport attributes. As such, if it can be successfully augmented as recommended, it should allow estimation of transport users' actual preferences, free of hypothetical biases.

* * *

References

[1] Abrantes, P. and M. Wardman, 2011, "Meta-Analysis of UK Values of Travel Time: An Update", *Transportation Research Part A*,45, 1–17.

[2] Berry, S., Levinsohn, J. and A. Pakes, 1995, Automobile Prices in Market Equilibrium, *Econometrica*, 63, 841–890.

[3] Bhouri, N., Aron, M. and J. Kauppilla, 2012, "Relevance of Travel Time Reliability Indicators: A Managed Lanes Case Study", *Procedia – Social and Behavioral Sciences*, 54, 450–459.

[4] Carrion, C. and D. Levinson, 2012, "Value of Travel Time Reliability: A Review of Current Evidence", *Transportation Research Part A*, 46, 720–741.

[5] Espino, R., Roman, C. and J. de Dios Ortuzar, 2006, "Analysing Demand for Suburban Trips: A Mixed RP/SP Model with Latent Variables and Interaction Effects", *Transportation*, 33, 241–261.

[6] Girotti, M. and R. Meade, 2015, U.S. Savings Banks' Demutualization and Depositor Welfare, November. Available at SSRN: `http://ssrn.com/abstract=2685235`.

[7] Hensher, D., 2006, "Integrating Accident and Travel Delay Externalities in an Urban Speed Reduction Context", *Transport Reviews*, 26(4), July, 521–534.

[8] Hensher, D., 2010, "Hypothetical Bias, Choice Experiments and Willingness to Pay", *Transportation Research Part B*, 44, 735–752.

[9] Hensher, D., 2011, "Valuation of Travel Time Savings", in de Palma, A., Lindsey, R. and E. Quinel (eds.), *Handbook of Transport Economics*, Edward Elgar.

[10] Hensher, D., 2015, "Customer Service Quality and Benchmarking in Public Transport Contracts", *International Journal of Quality Innovation*, 1(4).

[11] Hensher, D., Rose, J.. and W. Greene, 2015, *Applied Choice Analysis*, 2 ed., Cambridge University Press.

[12] Knittel, C. and K. Metaxoglou, 2014, Estimation of Random-Coecient Demand Models: Two Empiricists' Perspective, *Review of Economics and Statistics* , 96(1), March, 34–59.

[13] Li, Z., Hensher, D. and J. Rose, 2010, "Willingness to Pay for Travel Time Reliability in Passenger Transport: A Review and some New Empirical Evidence", *Transportation Research Part E*, 46, 384–403.

[14] Mazulla, G. and L. Eboli, 2006, "A Service Quality Experimental Measure for Public Transport", *European Transport*, 34, 42–53.

[15] McFadden, D., 1974, "The Measurement of Urban Travel Demand", *Journal of Public Economics*, 3, 303–328.

[16] McFadden, D., 2001, "Economic Choices", *American Economic Review*, 91, 351–78.

[17] McGirr, R., Bryant, J. and T. Hawkes, 2013, *Creating Regional Crime Statistics from Adminitrative Data*, Statistics New Zealand Working Paper No. 13–02.

[18] Ministry of Transport, 2015a, *Risk on the Road: Introduction and Mode Comparison*, August.

[19] Ministry of Transport, 2015b, *Risk on the Road: Drivers and their Passengers*, August.

[20] Ministry of Transport, 2015c, *Risk on the Road: Pedestrians, Cyclists and Motorcyclists*, August.

[21] Nevo, A., 2000, "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand", *Journal of Economics & Management Strategy*, 9(4), Winter, 513–548.

[22] New Zealand Police, 2016, *The Transformation of NZ Police Crime Statistics: New Measures and Trends*, February.

[23] NZTA, 2016, *Economic Evaluation Manual (Amendment 1).*

[24] OECD/ITF, 2010, *Improving Reliability on Surface Transport Networks.*

[25] Shires, J. and G. de Jong, 2009, "An International Meta-Analysis of Values of Travel Time Savings", *Evaluation and Programme Planning*, 32, 315–325.

[26] Small, K., 2012, "Valuation of Travel Time", *Economics of Transportation*, 1(1), 2–14.

[27] Small, K. and E. Verhoef, 2007, *The Economics of Urban Transportation*, Routledge, New York.

[28] Steimetz, S., 2008, “Defensive Driving and the External Costs of Accidents and Travel Delays”, *Transportation Research Part B*, 42, 703–724.

[29] Train, K., 2009, *Discrete Choice Methods with Simulation*, 2 ed., Cambridge University Press.

[30] van Oort, N., 2016, “Incorporating Enhanced Service Reliability of Public Transport in Cost-Benefit Analyses”, Public Transport, 8, 143–160.

[31] Vincent, M., 2008, *Measurement Valuation of Public Transport Reliability*, Land Transport New Zealand Research Report 339, Booz Allen Hamilton.

[32] Yansaneh, I., 2003, *Construction and Use of Sample Weights*, United Nations Secretariat Statistics Division.