WORKING PAPER

ITLS-WP-22-15

Light commercial vehicles destination choice: Understanding preferences relative to the number of stops and tour-based trip type

By

Camila Balbontin[a,b], David A. Hensher[a], Chinh Q. Ho[a]

[a] Institute of Transport and Logistics Studies (ITLS), The University of Sydney, Australia

[b] Engineering and Science Faculty
Universidad Adolfo Ibáñez, Santiago, Chile

August 2022

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-22-15

TITLE: Light commercial vehicles destination choice: Understanding preferences relative to the number of stops and tour-based trip type

ABSTRACT: Freight delivery modelling has made significant progress in the past few decades. In this study we propose to use an aggregate multi-step approach to gain a better understanding of the tour-based trips of light commercial vehicles in Sydney, Australia. The paper identifies preference differences in destination choice-making given by the number of stop and the tour-based trip type, defined by the total number of stops in the trip. The findings suggest that estimating a separate model for each number of stops provides a better understanding on how destination choices are made. Simulation results show how the probability of choosing a certain destination in the different tour-based trip types and number of stop changes due to variations in travel time and distance. Results show that light commercial vehicles are more sensitive to the generalised cost (defined by travel time and distance) in the first stop, and the sensitivity decreases as the trip is completed.

KEY WORDS: Light commercial vehicles; multi-step approach; destination choice; generalised cost; tour-based trips

AUTHORS: **Balbontin, Hensher, Ho**

ACKNOWLEDGEMENTS:

CONTACT: INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (H04)

The Australian Key Centre in Transport and Logistics Management

The University of Sydney NSW 2006 Australia

Telephone: +612 9114 1813

E-mail: business.itlsinfo@sydney.edu.au

Internet: http://sydney.edu.au/business/itls

DATE: August 2022

1 Introduction

Light commercial vehicles (LCV) tend to receive little attention in city and freight logistics studies even though they represent a significant traffic volume in cities, particularly in the last two years with the shift to online shopping and growing work opportunities for tradespeople who are in high demand for house renovations and repairs as more and more people work from home. Since the start of COVID-19, online shopping has increased dramatically in Australia which is expected to last well past this pandemic (Australia Post, 2021). The vast majority of online shopping deliveries are undertaken by LCVs, and thus their impact in cities has and will continue to increase in the coming years. This calls for a modelling framework that can account for LCV movement, in addition to passenger and heavy vehicle movement.

Tour-based stop choice models for LCV movement usually pool observations across multiple stops on the same service tour, effectively ignoring the sequence of stops in the decision-making process (see for example Ellison et al., 2017). This modelling approach is probably fine if the model implementation step uses a micro-simulation approach (cf. statistical approach) which relies on random draws to decide which locations are selected (i.e., the probability of a location being selected is either 1 or 0). A micro-simulation approach, however, requires a very large synthesis population to cover the entire probability distribution. Since the simulated choice is unique, one person has one choice at the time, be it mode, or time of day or location, etc. This will prolong the run time in application because the choice probability matrices become very large due to the large number of synthetic households or synthetic LCV workers, each with multiple choices (e.g., workplace, number of jobs/deliveries taking for the date, etc.). Regarding run time, the statistical approach has merits over the micro-simulation because the former requires only a few hundred synthetic households/LCV workers given to its ability to predict individual probabilities (ranging between 0 and 1) which when expanded, using the weight that each synthetic household/worker carries, would represent the choices made by the simulated population. The statistical approach, however, requires that all decisions regarding LCV movement, including the number of stops/deliveries, locations, and sequence of stops be modelled and properly connected to obtain internal consistencies (e.g., total number of stops visited across all service tours equals the number of deliveries required).

The paper proposes a multi-step modelling method to ensure internal consistency while speeding up running time. The multi-step approach uses a nested structure to describe a series of decisions LCV drivers make. These include the number of service tours to fulfil the job (e.g., deliver online shopping), the number of stops on each service tour and then the locations of these stops. Our proposed nested logit (NL) model uses the upper level to describe different tour types, defined by the number of stops chained onto a tour, and the lower level to decide *simultaneously* the location and sequence of stops visited before the vehicle is returned to the depot for the next tour or by the end of the day. The empirical model was estimated using data collected from 2007 to 2014 in Sydney, Australia through different sources. It contains information on the LCV industry and household location, LCV driver characteristics (e.g., age, gender, income), and destination characteristics such as time and distance from the current location, employment and populations density.

This paper is organised as follows. The next section provides a brief background of the light commercial vehicle literature and destination choice modelling. The following section explains the methodology used in this research for each of the models. Section 4 describes the data sources used in this research. The next section presents the model results and elasticities, and the final section discusses the most important findings.

2 Background

Freight modelling can be categorised into two streams: commodity-based and trip-based. Commodity-based modelling approaches freight movement from the commodity-based decisions in terms of the goods production and consumption and its relationship to economic characteristics, such as employment, industry size, and revenues (Hutchinson, 1974; Ogden, 1978; Wisetjindawat et al., 2006). By contrast, trip-based modelling, or its recent advancement – tour-based modelling, simulates vehicle movement as a chain of trips connecting multiple stops. Tour-based models that describe commercial vehicle movement can be further classified into two groups, based on the input data being disaggregate or aggregate.

Disaggregate models require data on each vehicle with different locations, destinations, vehicles and constraints; so they require a large amount of data with detailed information (Crainic et al., 2009; Ruan et al., 2012; Musolino et al., 2018; Thoen et al., 2020). Aggregate models estimate the average behaviour of vehicles, grouping them into categories. One aggregate model approach used is referred to as partial shares model which has two approaches: incremental growth or multi-step approach. Incremental growth studies consider that the vehicles decide where to stop first, and after each stop consider if they should go back to their base or continue (Hunt & Stefan, 2007; Wang & Holguín-Veras, 2008). The multi-step approach considers that the vehicle first decides how many stops will they do in their trip, and then decides the destinations of each stop (Nuzzolo & Comi, 2013; Ellison et al., 2017; Comi et al., 2021). Holguín-Veras et al. (2013) present a comprehensive overview of empirical findings and models that deal with urban freight tours.

Nuzzolo & Comi (2013) use the multi-step approach to understand tour-based visits carried out in Rome in 2008. They use delivery tour data collected on 500 truck drivers including medium and light goods vehicles. They define two models, one defining the number of stops and vehicle type (light or medium), and the second one the delivery location choices. The number of stops and vehicle type model include explanatory variables such as an accessibility index, distance between the origin and the study area (which is an inner zone of the city) and delivered quantity. The delivery location choices model was separated into the first delivery location model and the next delivery location zone model. The explanatory variables include retail employees in the destination area, an accessibility index (which takes into account accessibility indicators such as area pricing, route constraints or vehicle type constraints), distance, delivery share, and a memory variable. The memory variable represents the history of the tour, i.e., the ratio between the distance to be covered to reach the next delivery location and the current distance covered. Their results show that there are important differences between the first stop and subsequent stops, with the memory variable being statistically significant and negative, suggesting that the systematic utility of choosing a destination is a cost function of the distance from the current stop location. Comi et al. (2021) use an aggregate approach to estimate the tour-based visits of light goods vehicles in the Veneto Region, Italy. Their approach is similar to the one described by Nuzzolo & Comi (2013), where they separate the first delivery zone and the next zones. They used data collected using global position system (GPS) which enables them to obtain true freight origin- destination matrices.

Wang & Holguín-Veras (2008) use an incremental growth approach, also known as a stop-go model, to describe the probability of terminating the trip and returning to the origin. They use a synthetic dataset that contains a set of tours. They separate their modelling into the destination choice and the tour termination process. Their results show that the termination model is a function of a memory variable that represents the amount of cargo the truck has delivered up to that point in the termination model, and the distance to the origin. This study provides insights on the variables that should be considered in modelling freight movement. Hunt & Stefan (2007) use an incremental growth approach to understand tour-based movements in Calgary, Canada in 2001. Their study includes light vehicles, heavier single unit and multi-unit configurations. They include explanatory variables such as travel time, primary purpose, employment, residential land use, have a statistically significant influence in the next stop destination model.

Boarnet et al. (2017) study freight movements at an aggregate level in Los Angeles in 2005. They estimate a regression model to understand freight activity as a function of geographic characteristics, employment, and accessibility. Their results show that employment is an important driver of freight activity.

Ellison et al. (2017) study light commercial vehicle destination choices in Sydney. They use a multi-step approach to modelling destinations visited, but they do not estimate a model for the number of stops visited or the sequence of these visits. That is, the utility function for a location to be chosen is the same, with the stop sequence not modelled per se but simulated by random draws. Their findings suggest that occupation and industry are very important explanatory variables in light commercial vehicle trip making.

The aforementioned studies provide relevant guidelines as to what are the most important variables to include. In the current paper, we propose a nested structure and estimate a separate model for each stop number. The models from the upper levels are connected to the ones in the lower levels through a logsum parameter, an index of expected maximum utility. This structure directly accounts for the cumulative effect of the number of stops and the distance travelled on tour formation and destination choices within a utility-maximisation framework that is internally consistent and quick in run time. This is the main contribution of this paper which, to the best of our knowledge, has not been done before in the LCV and freight modelling literature.

3 Methodology

The aim of this paper is to estimate an operational LCV model that can be implemented within an integrated transport and land use modelling framework, called MetroScan (see, Hensher et al., 2020) to account for the competing demands for roads generated by moving passengers, services, and freight. The contribution of the method used in this research and main focus of this paper is on the LCV preferences' variations across different stops and different tour types. To this end, we developed a tour-based nested logit model structure for the movement of LCVs based on the Sydney Household Travel Survey (HTS) 7-year pooled dataset from which all movement of LCVs registered for business (cf. private) were filtered and restructured into a tour-based dataset (see Ho & Mulley, 2013 for more details). Our model structure reflects the interdependent decisions observed for this sector, namely tour generation (i.e., number of tours generated per day), tour type (i.e., number of stops on each tour), destination locations and stop sequence (i.e., which locations to visit in what order). The tour generation sub-model is structured to as a NL model with the number of levels equal to the maximum number of tours observed for each LCV driver. The tour type sub-model was also structured to have a NL form with five nests, each representing the number of stops chained into that tour. The two sub-models were linked with the destination and stop sequence choice model through the maximum expected utility concept (aka. logsum see, e.g., Hensher et al., 2015). The latter is the focus of this paper.

The destination and stop sequence choice model uses a travel zone (TZ) defined by Transport for New South Wales (TfNSW) as the smallest spatial unit of geography.[1] There are over 2,000 TZs within the Greater Sydney Metropolitan Area (GSMA). Thus, using TZs as elementary choices of the NL model results in a very large choice set, presenting challenges for model estimation and application. Thus, the technique we used to go around this large choice set issue involved separating, *econometrically*, the destination/stop sequence decision into two steps: (1) the driver first decides which Statistical Area Level 3 (SA3, a relatively large spatial unit of geography defined by the relevant authorities) to visit and then (2) decides which TZ to visit within the selected SA3. The GSMA is divided into 58 SA3s, each contains multiple TZs (the maximum TZs within a SA3 is 123). Hence, the choice set reduces from more than 2000 alternatives (TZs) to 58 for the choice of SA3 and a maximum of 123 for the choice of TZ. This technique allows us to estimate a destination and stop sequence choice model simultaneously, without imposing arbitrary constraints on the choice set. That is, all SA3 and TZs are available for all LCVs.

We estimated the destination and stop sequence choice model using the *multi-step approach* with *partial share models* (see Section 2). The multi-step approach model structure is illustrated in Figure 1. The LCV driver first chooses how many stops to visit per tour (i.e., tour type). Given a positive number of stops, they will then need to decide what stops (SA3 location, followed by TZ location) to visit first, second, third, etc. until all stops are visited. For example, if the tour includes two stops, they need to decide, *in sequential order*, the SA3 and the TZ of the first stop, then the SA3 and TZ for the second stop before returning to their home or the depot for the next tour. The decision levels within the same branch (i.e., TZ stop t → SA3 stop t → TZ stop t-1 → SA3 stop t-1 → …) are connected via a logsum which is obtained from the bottom levels to feed into the top levels. This logsum represents the utilities of the bottom levels which are required to define the utilities of the top levels. For example, the TZ model for stop 3 in branch 4 (i.e., with a total number of 4 stops), will influence the top-level model, which is the SA3 model for stop 3 in branch 4 stops. Similarly, the SA3 model for stop 3 in branch 4 will influence the TZ model for stop 2 in branch 4.

The last choice, going back to their home/depot, always has a probability of 1, owing to the assumption that all LCVs return to the home/depot by the end of the day. Therefore, for model estimation, the trips going back to the home/depot will be excluded; however, in implementation, this so-called *empty trip* is added to the LCV demand matrix, using the last TZ as the origin and the driver's home/depot as the destination, before passing on to the network assignment. We estimated separate models for different tour types (i.e., branch). This approach will allow us to gain a better understanding of how the different attributes are more or less important in the different tour types (i.e., total number of stops) and throughout the trajectory (i.e., between stops in the same tour). The models for branch with 5 or more stops represent the longest sequential estimation, since it has the longest decision-process as was presented in Figure 1, with 10 levels – hence 10 models excluding the number of stops model. The models for each branch will be estimated sequentially.

The next subsections provide more details of the model formulation for the TZ destination choice model, the SA3 destination choice models, and the number of stops choice. The last two subsections (3.4 and 3.5) detail the approach used to define the generalised cost used in the models, which represents the time and distance between the origin and destinations, including: (1) generalised cost between the origin and the possible SA3s; (2) generalised cost between the chosen SA3 and the

[1] https://opendata.transport.nsw.gov.au/dataset/travel-zones-2016

possible TZs within the SA3, referred to as normalised generalised cost; and (3) generalised cost between the origin and the driver's home/depot.

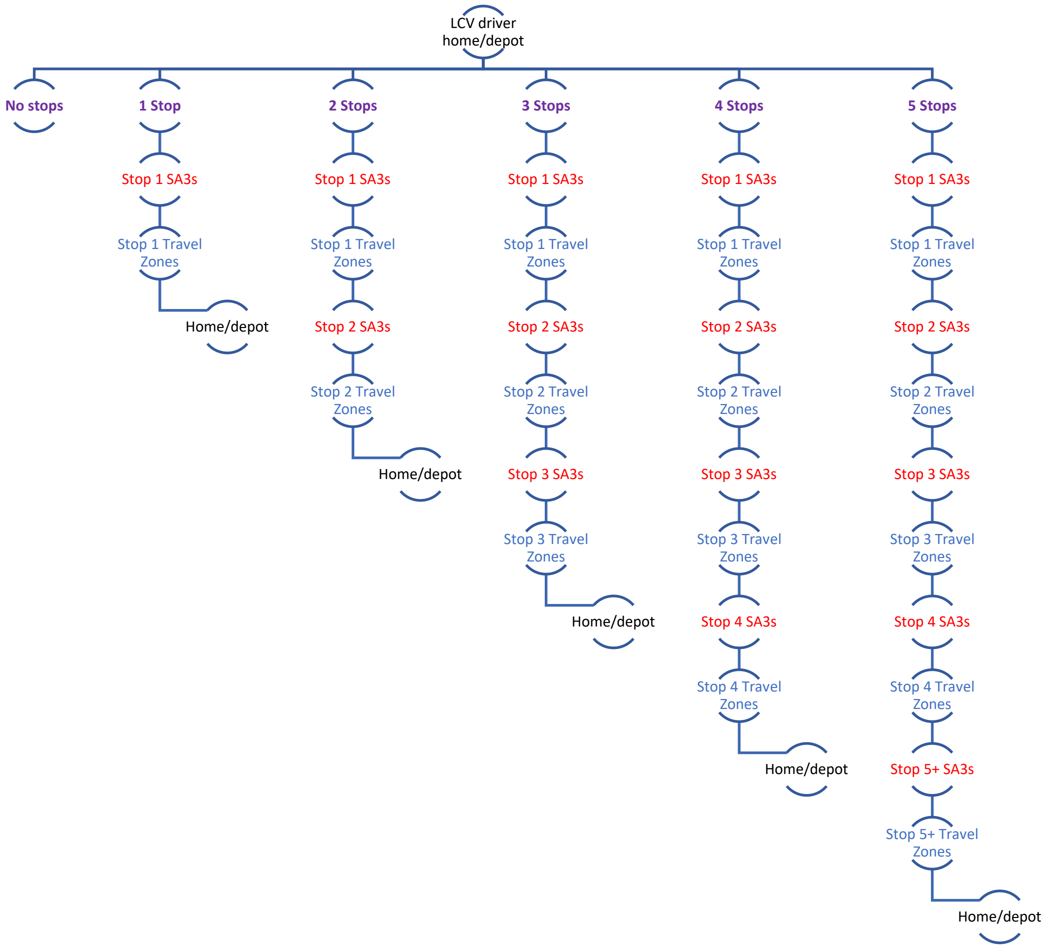


Figure 1: Nested structure used for sequential model estimation

3.1 TZ Models

The first model to be estimated in each branch is the one in the lowest level, i.e., the TZ model for the last stop *S* (in branch *S,* with *S* stops). This model is estimated using the observation of the last stop *S,* for LCVs that did a tour with *S* stops (i.e., belong to branch *S*). The utility function for destination TZ_i given the LCV driver already choose the SA3 destination, is as follows:

$$U^{S_S Br_S}_{O,DTZ_i^{SS}|DSA3^{SS}} = \left(\beta^{SS}_{GC} + \sum_{j\in ind} \beta^{SS}_{ind_j} \cdot \mathrm{ind}_j + \beta^{SS}_{sex} \cdot \mathrm{sex} + \beta^{SS}_{age} \cdot \ln(\mathrm{age}) + \beta^{SS}_{hinc} \cdot \ln(\mathrm{hinc}) \right) \cdot \mathrm{GenCost}^{n}_{DSA3^{SS},DTZ_i^{SS}}$$
$$+ \sum_{m\in \mathrm{ind}} \beta^{SS}_{Empl_m} \cdot \ln\left(\mathrm{Empl}_{m,DTZ_i^{SS}}\right) + \beta^{SS}_{Pop} \cdot \ln\left(\mathrm{Population}_{DTZ_i^{SS}}\right)$$

(1)

where ind_j is a dummy variable equal to 1 if the LCV driver is in industry *j*, 0 otherwise; sex is the gender of the driver; age is the age of the LCV driver; hinc the annual household income of the LCV driver (in AUD$); $\mathrm{GenCost}^{n}_{DSA3^{SS},DTZ_i^{SS}}$ represents the normalised generalised cost between the chosen SA3 for stop *S* and the TZ *i*; $\mathrm{Empl}_{m,DTZ_i^{SS}}$ represents the employment (number of employed people) by industry *m* in the destination TZ *i*; and $\mathrm{Population}_{DTZ_i^{SS}}$ represents the estimated resident population in TZ *i*. The β s are the parameters associated to each attribute. The attribute normalised generalised cost is associated with several parameters: β_{GC}^{SS} by itself, β_{sex}^{SS} interacting with the driver's gender, β_{age}^{SS} interacting with driver's age and β_{hinc}^{SS} interacting with driver's income.

The TZ utility formulation for the rest of the stops (i.e., that are not the last one) is the same as the one presented in Equation (4) but it includes the logsum parameter for the bottom SA3 model (refer to Figure 1). That is, the logsum parameter for TZ stop *t* model is calculated using the SA3 model for stop *t*+1, $\mathrm{lnSum}_{DSA3^{\mathrm{S(t+1)}}}^{St}$. The utility expression for the TZ model for stop *t* (*t < S)* can be written as:

$$U_{DTZ_i^{\mathrm{St}}|DSA3^{\mathrm{St}}}^{S_tBr_S} = \left(\beta_{GC}^{St} + \sum_{j \in ind} \beta_j^{St} \cdot \mathrm{ind}_j + \beta_{sex}^{St} \cdot \mathrm{sex} + \beta_{age}^{St} \cdot \ln(\mathrm{age}) + \beta_{hinc}^{St} \cdot \ln(\mathrm{hinc})\right) \cdot \mathrm{GenCost}^{\mathrm{n}}_{DSA3^{\mathrm{St}}, DTZ_i^{\mathrm{St}}}$$
$$+ \sum_{m \in \mathrm{Ind}} \beta_{Empl_m}^{St} \cdot \ln\left(\mathrm{Empl}_{\mathrm{m}, DTZ_i^{\mathrm{St}}}\right) + \beta_{Pop}^{\mathrm{St}} \cdot \ln\left(\mathrm{Population}_{DTZ_i^{\mathrm{St}}}\right) + \beta_{\mathrm{lnsum}}^{\mathrm{St}} \cdot \mathrm{lnSum}_{DSA3^{\mathrm{S(t+1)}}}^{St} \tag{2}$$

where $\beta_{\mathrm{lnsum}}^{\mathrm{St}}$ represents the parameter associated with the logsum of the SA3 model for the next stop *t+1,* and it has to be between 0 and 1 to comply with the underlying assumption of expected utility maximisation*.* If it is equal to 0, then the results would be suggesting that the SA3 for *t+1* stop model does not have an influence on the current stop *t* TZ model.

3.2 SA3 Models

The SA3 model for stop *t* (*t ≤ S)* branch *S* will be estimated sequentially after the TZ stop *t* branch *S* model, as it is related to the bottom model through the logsum of the TZ model, $\mathrm{lnSum}_{DTZ^{\mathrm{St}}}^{St}$. The utility function for SA3 *i* in stop *t* Branch *S,* given that the previous TZ stop *t-1* has already been chosen, will be defined as:

$$U_{DSA3_i^{\mathrm{St}}|DTZ^{\mathrm{S(t-1)}}}^{S_tBr_S} = \left(\beta_{GC}^{St} + \sum_{j \in ind} \beta_j^{St} \cdot \mathrm{ind}_j + \beta_{sex}^{St} \cdot \mathrm{sex} + \beta_{age}^{St} \cdot \ln(\mathrm{age}) + \beta_{hinc}^{St} \cdot \ln(\mathrm{hinc})\right) \cdot \ln\left(\mathrm{GenCost}_{DTZ^{\mathrm{S(t-1)}}, DSA3_i^{\mathrm{St}}}\right)$$
$$+ \beta_{GCHH}^{St} \cdot \ln\left(\mathrm{GenCostHH}_{DSA3_i^{\mathrm{St}}, DTZ_{HH}}\right) + \beta_{CBD}^{\mathrm{St}} \cdot \mathrm{CBD}_{DSA3_i^{\mathrm{St}}} + \beta_{\mathrm{lnsum}}^{\mathrm{St}} \cdot \mathrm{lnSum}_{DTZ^{\mathrm{St}}}^{St} \tag{3}$$

where $\mathrm{GenCost}_{DTZ^{\mathrm{S(t-1)}}, DSA3_i^{\mathrm{St}}}$ represents the generalised cost between the TZ chosen in the previous stop *t-1* and SA3 *i*; $\mathrm{GenCostHH}_{DSA3_i^{\mathrm{St}}, DTZ_{HH}}$ represents the generalised cost between SA3 *i* and the TZ driver's home/depot (i.e., representing the time and distance that would take the driver to go back to their origin); and $\mathrm{CBD}_{DSA3_i^{\mathrm{St}}}$ is a dummy variable equal to 1 if SA3 *i* is located in the Central Business District (CBD) area, 0 otherwise. The β s are the parameters associated to each attribute and, similar to the TZ models, the generalised cost interacts with the LCV driver characteristics (industry, age, gender, income).

3.3 Tour Type Model

The tour type model, describing the number of stops an LCV driver visits on each tour, is the last to be estimated because it includes the logsum of the destination and stop sequence choice model. The utility function for branch *b* (b between 1 to 5+ stops) will be defined as follows:

$$\mathrm{U}^{Br_b} = ASC_b + \beta_{CBD_b} \cdot CBD_{HH} + \sum_{j \in ind} \beta_{j_b} \cdot \mathrm{ind}_j + \beta_{sex_b} \cdot sex + \beta_{age_b} \cdot \ln(age) + \beta_{hinc_b} \cdot \ln(\mathrm{hinc})$$
$$+ \beta_{weekend_b} \cdot weekend + \beta_{\mathrm{lnsum_b}} \cdot \mathrm{lnSum}_{DSA3^{\mathrm{S1}}}^{Br_b} \tag{4}$$

where CBD_{HH} is a dummy variable equal to 1 if the LCV driver home/depot is located in the CBD area, 0 otherwise; $weekend$ is a dummy variable equal to 1 if the tour is completed on a weekend day, 0 otherwise; $\ln\text{Sum}_{DSA3^{S1}}^{Br_b}$ represents the logsum of the first stop of branch *b*; and the β s are the associated parameter estimates.

3.4 Generalised Cost and Normalised Generalised Cost

The nested structure considers the LCV first choosing the SA3 to stop (SA3 destination) and, given that SA3, they choose the TZ within that SA3. For the SA3 decision, we assume LCV drivers consider the distance between their origin and the possible SA3 destinations $d_{OTZ,DSA3}$ - which is calculated as the average distance between their origin (travel zone) and all the travel zones within each SA3 destination. For the TZ decision, given that LCV drivers already chose the SA3 destination, we assume they evaluate a "normalised" distance $d^n{}_{DSA3,DTZ}$ which is the distance between their chosen SA3 destination and the possible TZ destinations, calculated as:

$$d^n{}_{DSA3,DTZ} = d_{OTZ,DTZ} - d_{OTZ,DSA3} \qquad \textbf{(5)}$$

where $d_{OTZ,DTZ}$ is the distance between the origin TZ to destination TZ (given by the OD matrix which is represented by travel zones). We calculate the equivalent for time, and calculate a generalised cost as follows:

$$\begin{aligned}\text{GenCost}_{OTZ,DTZ} &= t_{OTZ,DTZ} + \frac{\text{VOC}}{\text{VTTS}} \cdot d_{OTZ,DTZ} = \text{TimeAM}_{OTZ,DTZ} + \frac{\text{VOC}}{\text{VTTS}} \cdot \text{DistAM}_{OTZ,DTZ} \\ &= \text{TimeAM}_{OTZ,DTZ} + \frac{37 \text{ cents/km}}{58.33 \text{ cents/min}} \cdot \text{DistAM}_{OTZ,DTZ}\end{aligned} \qquad (6)$$

where VOC is the value of operating costs which is equal to 37cents/km (ATAP, 2016). The VTTS used is derived from several non-commuting work-related studies (Mackie et al., 2003; Ellison et al., 2017; Batley, 2015), which is 35 \$/hour (58.33 cents/minute).

The normalised generalised cost which are used in the TZ decisions assuming the LCV driver has already chosen the SA3, are as follows:

$$\begin{aligned}\text{GenCost}^n{}_{DSA3,DTZ} &= t^n{}_{DSA3,DTZ} + \frac{\text{VOC}}{\text{VTTS}} \cdot d^n{}_{DSA3,DTZ} \\ &= \left(t_{OTZ,DTZ} - t_{OTZ,DSA3}\right) + \frac{\text{VOC}}{\text{VTTS}} \cdot \left(d_{OTZ,DTZ} - d_{OTZ,DSA3}\right)\end{aligned} \qquad (7)$$

A graphical representation of these normalised distances calculation is presented in Figure 2, which is represented by the black line. This is the difference between: the distance between travel zones (light blue line) – which is obtained from the OD matrix; and the distance between the travel zone and the SA3 centre (green line).

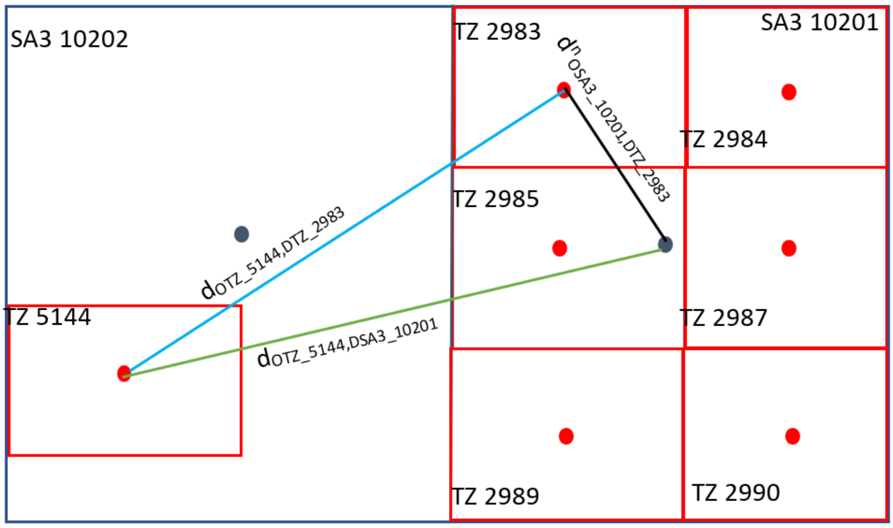


Figure 2: Graphical representation of normalised distance between SA3 origin and TZ destination

As mentioned above, the centre of each SA3 is calculated as the average distance between all the travel zones within that SA3 (not only the ones chosen in the dataset, but all available according to the OD matrix). An example is shown in Figure 3, where the average between all the light blue distances (between TZs) is equal to the green distance (distance between a TZ and the centre of the SA3).

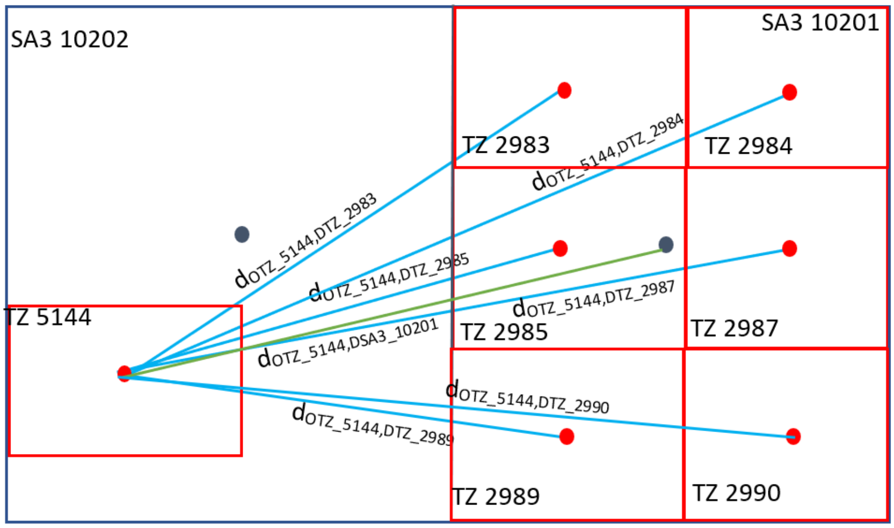


Figure 3: Graphical representation of the centre of each SA3

It is important to note that the normalised distance (represented by the black line in Figure 2) can be negative, when the destination travel zone is closer to the travel zone origin, relative to the SA3 centre; or positive, when the destination travel zone is more distant to the origin than the SA3 centre. An example is presented in Figure 4. The TZ 2984 is more distant to the origin (TZ 5144) than the SA3 10201 centre, so the normalised distance for TZ 2984 from the SA3 10201 centre is positive and represented by the dotted black line. Contrarily, TZ 2985 is closer to the origin (TZ 5144) than the SA3 10201 centre, so the normalised distance for TZ 2985 from the SA3 10201 centre is negative and represented by the solid black line.

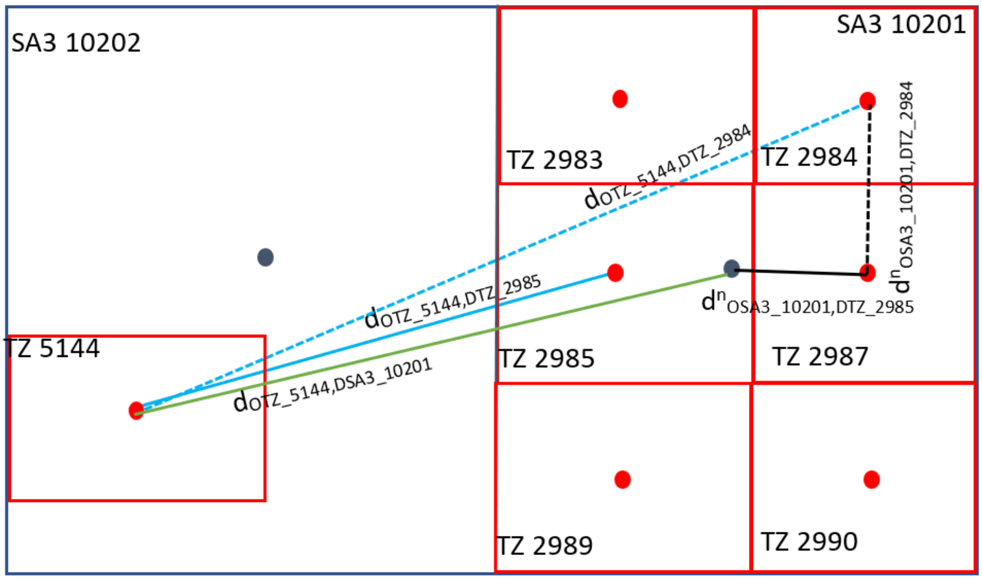


Figure 4: Graphical representation of positive (dotted lines) and negative (solid lines) normalised distances between a SA3 centre and a travel zone

A graphical representation of an example of the normalised distances for all travel zones within a SA3 is shown in Figure 5.

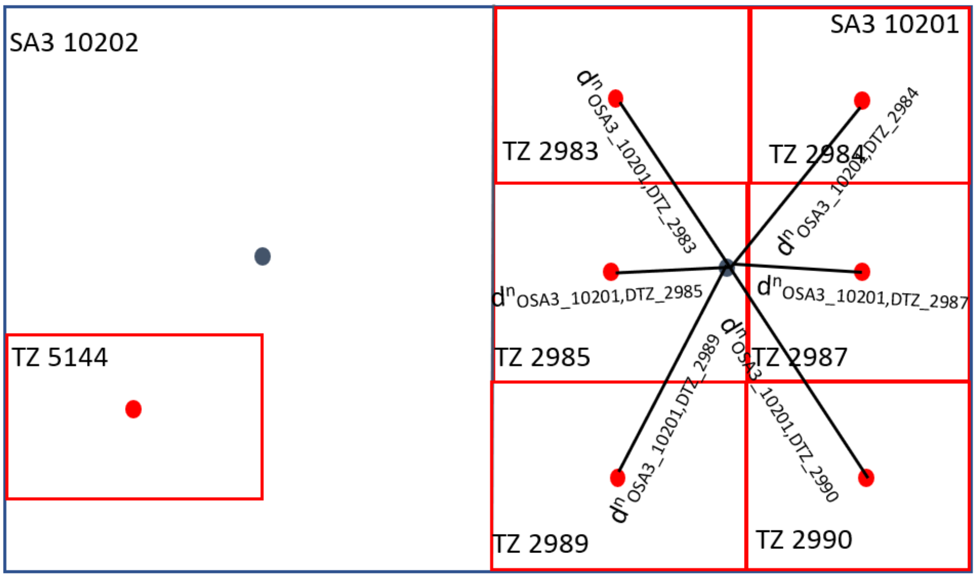


Figure 5: Graphical representation of normalised distances within an SA3

3.5 Generalised Cost to Home/Depot

In terms of destination choice, it seems plausible that LCVs consider the distance and time back to their origin, whether it is the LCV driver's home or depot, when deciding where to go next. This generalised cost to home/depot was included in the SA3 models as presented in subsection 3.2. To be able to include the log generalised cost back to the household, *lGenCostHH*, we need to verify that its correlation with the generalised cost between the current location and the SA3 alternatives, *lGenCost*, does not cause a severe multicollinearity problem, using a rule of thumb value of 0.8 for partial correlation. That is, if the correlation between *lGenCost* and *lGenCostHH* is higher than 0.8 then we only include *lGenCost* in the model.

The correlation results are presented in Table 1. The generalised costs for the SA3 first stop are obviously highly correlated since they represent the same two points but in opposite direction (going to the SA3 or coming back to the house which is the starting point), so the distance is the same, but the time slightly changes. In all branches, the generalised costs are highly correlated for stop 2 as well, so the generalised cost back household are not included in any of the stop 1 or 2 SA3 models.

Table 1: Correlation between *lGenCost* and *lGenCostHH*

Correlation *lGenCost* and *lGenCostHH* in SA3 models	Branch stop 1	Branch stop 2	Branch stop 3	Branch stop 4	Branch stop 5
Stop 1	0.966	0.956	0.962	0.961	0.958
Stop 2	-	0.840	0.851	0.859	0.880
Stop 3	-	-	0.788	0.787	0.796
Stop 4	-	-	-	0.791	0.776
Stop 5	-	-	-	-	0.790

4 Data

The data used in this study was obtained pooling four sources of data:

1. **LCV movement:** These data include LCV tours constructed from the 7-year pooled Sydney Household Travel Survey data (2007 to 2014). This is the base data for the number of trips and destination choice modelling (TZ and SA3).
2. **Employment and population data**: Includes employment and populations projections per industry class.
3. **Distance and time OD matrix:** These data contain all the information for times and distances between each TZ pairs.

The TZ and SA3 destination choice models, as well as the number of trips model, were estimated using the pooled freight trips data, employment and population data, and the distance and time OD matrix (used to calculate the generalised costs). These pooled data are referred to as the destination choice data, which contains information for 1,302 tour-based trips; and the expanded version that considers each stop separately contains 3,649 observations for all tour-based trips. The freight trips data also contain information on the age, gender and household income of the LCV drivers (in '000$AUD), as well as their occupation and industry.

The occupation of LCVs driver is presented in Figure 6, where 45.8% of these data represents people with occupation *technicians and trade workers,* 16.9% represents *managers,* 11.6% represents *professionals,* etc. Even though almost half of the observations are represented by *technicians and trade workers*, the other half of work-related tours corresponds to different occupations.

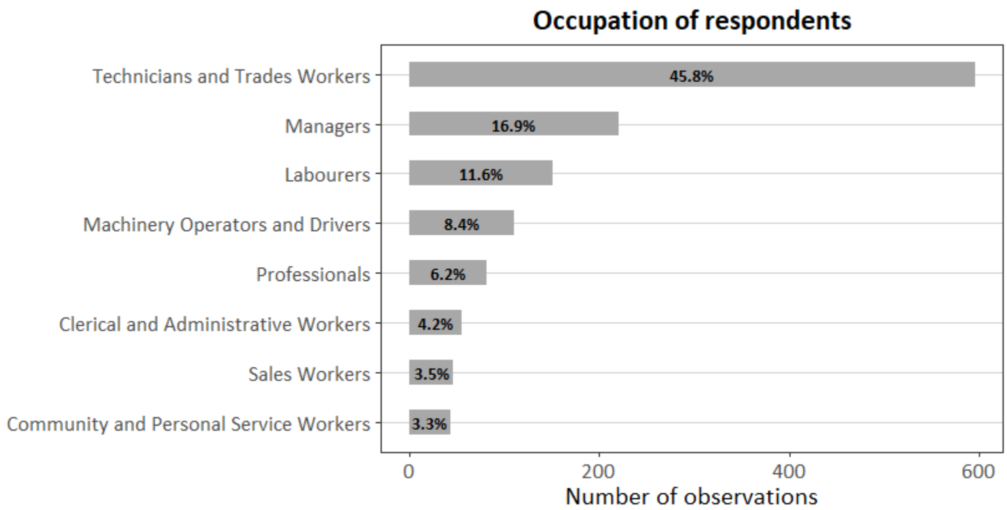
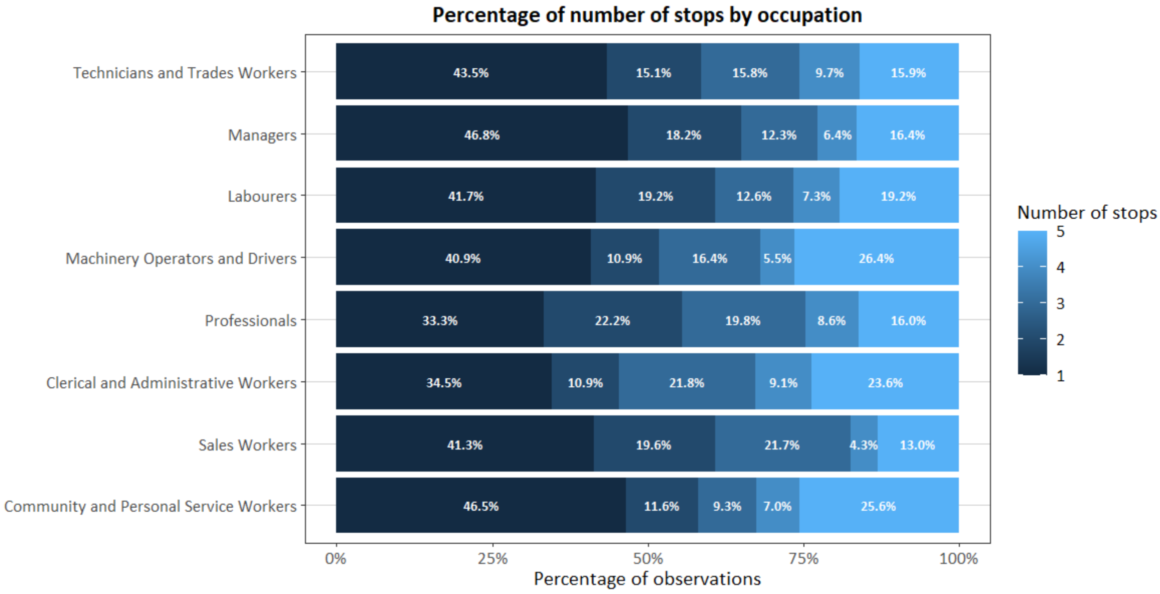


Figure 6: Occupation of LCV drivers

The percentage of tour-based trip types by occupation is presented in Figure 7. The distribution across occupations is relatively similar, that is, the most frequent branch is the 1-stop branch, and the lowest is the 4-stop branch. The 2, 3 and 5-stop branches have different percentages across occupations. The 5-stop branch is close or over 25% in machinery operators and drivers, clerical and administrative workers, and community and personal service workers. For sales workers, the second most common branch is the 3-stop branch. Therefore, it is important to note that the number of stops model will be representing the technicians and trades workers shares, where the 2, 3 and 5-stop branches are almost equally frequent. It is important to note that even though the branch with lowest number of observations is the 4-stops branch, it still has over 100 valid tour-based trips, which is enough to estimate the TZ and SA3 models. All the other tour-based trip types will be estimated with over 200 observations.



Note: Number of stops excludes the last household stop

Figure 7: Tour-based trip types by occupation

Table 2 presents the LCV occupation by industry. Most of the technicians and trade workers are in the construction sector, while the majority of sales workers and managers are in the wholesale and retail trade industry, and the majority of the machinery operators and drivers work in transport and warehousing. As expected, occupations are highly related to the industry type, which is what will be included in the models.

Table 2: LCV occupation by industry

Industry/Occupation	Clerical and Administrative Workers	Community and Personal Service Workers	Labourers	Machinery Operators and Drivers	Managers	Professionals	Sales Workers	Technicians and Trades Workers	Total
Accommodation and food services		19%	2%		4%		4%	1%	2%
Agriculture and mining	4%		8%	1%	5%			1%	2%
Construction	2%		28%	13%	16%	1%		58%	34%
Education and training		5%	5%		0%	17%		1%	2%
Financial and insurance services	4%				2%	2%			1%
Health care and social assistance	5%	30%	3%		1%	14%		1%	3%
Manufacturing	22%		9%	20%	19%	10%	9%	14%	14%
Media, professional, scientific and ICT services	11%		1%	1%	4%	23%	9%	4%	5%
Public administration	15%	16%	3%	3%	1%	12%		3%	4%
Real estate, administration and other services	15%	30%	30%	4%	9%	14%	9%	12%	14%
Transport and warehousing	11%		5%	40%	5%	5%		1%	6%
Wholesale and retail trade	13%		5%	19%	34%	1%	70%	5%	13%

General descriptives of the TZ and SA3 destination choice datasets and the number of stops dataset are presented in Table 3. Note that the average normalised generalised cost in the TZ destination choice data is negative, which means that LCVs tend to choose a TZ that is closer to their origin than the centre of the chosen SA3. The average distance to the SA3 chosen destination is approximately 17.6 kms, while the travel time by car is of 31.6 minutes. The average age of LCV drivers in the destination choice data is almost 42 years, 12% of the drivers are female and their average household annual income is of AUD$110k.

Table 3: Sample descriptive statistics

Variables	Mean (std dev)	Data level
Number of stops data		
Household in the CBD area (1,0)	17%	Stops
Weekend (1,0)	8%	Stops
Destination choice data		
Distance to household (kms)	21.739 (20.29)	SA3
Distance to SA3 (kms)	17.597 (18.01)	SA3
Generalised cost	42.733 (38.63)	SA3
Generalised cost to household	48.969 (39.45)	SA3
If the destination SA3 is in the CBD area (1,0)	27%	SA3
Travel time by car to household (minutes)	35.181 (27.34)	SA3
Travel time by car to SA3 (minutes)	31.572 (28.04)	SA3
Accommodation and food services industry employment in TZ destination	136.189 (201.83)	TZ
Agriculture and mining industry employment in TZ destination	18.698 (57.32)	TZ
Construction industry employment in TZ destination	179.687 (271.72)	TZ
Education and training industry employment in TZ destination	91.805 (206.94)	TZ
Estimated resident population by destination TZ	1,876.924 (1,371.35)	TZ
Financial and insurance services industry employment in TZ destination	96.939 (421.05)	TZ
Health care and social assistance industry employment in TZ destination	214.047 (628.55)	TZ
Manufacturing industry employment in TZ destination	313.500 (624.63)	TZ
Media, professional, scientific and ICT services industry employment in TZ destination	206.977 (494.03)	TZ
Normalised generalised cost	-7.153 (13.28)	TZ
Public administration industry employment in TZ destination	130.531 (385.04)	TZ
Real estate, administration and other services industry employment in TZ destination	238.518 (321.78)	TZ

Variables	Mean (std dev)	Data level
Transport and warehousing industry employment in TZ destination	178.037 (793.10)	TZ
Wholesale and retail trade industry employment in TZ destination	366.386 (541.91)	TZ
Age of driver	41.830 (13.54)	LCV driver
Female driver (1,0)	12%	LCV driver
Household income of driver in '000$AUD	110.246 (54.45)	LCV driver

5 Results Multi-Step Approach

The TZ model results for each branch are presented in Table 4 and Table 5, and the SA3 models results are presented in Table 6 and Table 7. The overall goodness of fit of the models presented at the bottom of the tables shows that the models are statistically superior to the restricted version, suggesting the explanatory variables are statistically significant when explaining the destination choices. The results show that the TZ decision for the current stop (t) does have a statistically significant effect in the current stop (t) SA3 decision, as the logsum representing the bottom TZ model was statistically significant in all SA3 models. This result was expected since the TZ options within a SA3 certainly have an influence on the SA3 chosen for each stop. On the contrary, results show that the relationship between stops was not statistically significant as the logsum representing the bottom SA3 decision was not statistically significant in any of TZ models. That is, the SA3 decision for the next stop (t+1) does not have a significant effect in the current stop (t) TZ decision. This suggests that LCV drivers may not consider the convenience of the next stop, measured by the maximum expected utility of visiting that stop, when deciding where to stop next.

It is interesting to note that there are important differences for the explanatory variables that are statistically significant in each stop. For instance, the TZ model results for branch 5 stop 5 show that LCV drivers that work in accommodation and food services industry are less sensitive to the normalised generalised cost (nGenCost), followed by LCV drivers that are in the construction industry, compared to the rest of the industries. They also show that as the LCV driver is older, they are less sensitive to the nGenCost in stop 5, and female drivers are more sensitive to the nGenCost. In the TZ model results for branch 5 stop 4, the only LCV driver characteristic that was statistically significant was the manufacturing industry, and it shows that LCV drivers that work in that industry are less sensitive to the nGenCost in stop 4. The TZ model results for branch 5 stop 3 show that LCV drivers in the financial and insurance services industry are more sensitive to the nGenCost in this stop. The only attribute that was statistically significant in all the models for branch 5 was the nGenCost and the employment of the driver in the transport and warehousing industry, which suggests that LCV drivers who are doing a 5-stop tour are more likely to stop in TZ that have higher employment in transport and warehousing industry. Similarly, the results for the rest of the branches also vary across stops with the nGenCost being the only statistically significant attribute in all models.

Table 4: TZ Model Results for Branch 5 and 4 – mean (t value)

Attribute	Branch 5					Branch 4			
	Stop 1	**Stop 2**	**Stop 3**	**Stop 4**	**Stop 5**	**Stop 1**	**Stop 2**	**Stop 3**	**Stop 4**
Normalised generalised cost from the current SA3 centroid to alternative TZ destination	-0.059 (7.39)	-0.074 (8.58)	-0.077 (8.49)	-0.101 (8.99)	-0.308 (4.97)	-0.074 (6.29)	-0.068 (5.77)	-0.072 (5.39)	-0.083 (6.38)
Industry type manufacturing (1,0)				0.069 (2.63)					
Industry type construction (1,0)					0.024 (1.85)				
Industry type accommodation and food services (1,0)					0.073 (1.88)				
Industry type financial and insurance services (1,0)			-0.832 (2.16)						
Age of driver					0.057 (3.47)				
Female driver (1,0)	-0.056 (1.97)				-0.065 (3.16)				
Ln(Estimated resident population by destination TZ)		-0.057 (1.92)				-0.179 (4.17)			
Agriculture and mining industry employment in TZ destination							0.143 (1.77)		
Manufacturing industry employment in TZ destination							0.207 (3.56)		
Construction industry employment in TZ destination	0.247 (3.12)								
Wholesale and retail trade industry employment in TZ destination					0.086 (2.05)				0.265 (3.49)
Accommodation and food services industry employment in TZ destination			0.262 (5.28)						
Transport and warehousing industry employment in TZ destination	0.155 (2.77)	0.102 (1.85)	0.197 (4.25)	0.149 (3.14)	0.103 (3.20)				
Media, professional, scientific and ICT services industry employment in TZ destination								0.284 (2.10)	
Public administration industry employment in TZ destination				0.151 (3.83)	0.082 (3.21)				0.115 (1.73)
Health care and social assistance industry employment in TZ destination	0.114 (2.45)					0.139 (2.15)			-0.201 (2.73)
Real estate, administration and other services industry employment in TZ destination		0.334 (4.54)			0.123 (1.95)			0.297 (2.01)	
Number of estimated parameters	***5***	***4***	***4***	***4***	***9***	***3***	***3***	***3***	***4***
Sample size	***232***	***232***	***232***	***232***	***724***	***106***	***106***	***106***	***106***
Restricted log likelihood	***-1,116.43***	***-1,116.43***	***-1,116.43***	***-1,116.43***	***-3,484.02***	***-510.09***	***-510.09***	***-510.09***	***-510.09***
Final log likelihood	***-787.19***	***-781.70***	***-784.66***	***-778.73***	***-2,376.81***	***-375.95***	***-368.92***	***-360.09***	***-355.34***
AIC/n	***6.83***	***6.77***	***6.80***	***6.75***	***6.59***	***7.15***	***7.02***	***6.85***	***6.78***

Table 5: TZ Model Results for Branch 3, 2 and 1 – mean (t value)

Attribute	Branch 3			Branch 2		Branch 1
	Stop 1	Stop 2	Stop 3	Stop 1	Stop 2	Stop 1
Normalised generalised cost from current SA3 centroid to alternative TZ destination	-0.037 (4.85)	-0.173 (2.83)	-0.072 (7.78)	-0.058 (7.94)	-0.183 (2.66)	-0.046 (10.41)
Industry type agriculture and mining (1,0)						-0.066 (2.44)
Industry type accommodation and food services (1,0)	0.203 (3.61)					
Industry type health care and social assistance (1,0)					-0.153 (2.22)	
Female driver (1,0)	-0.069 (2.18)					
Ln(Household income of LCV driver in '000$AUD)		0.027 (1.95)			0.031 (2.07)	
Agriculture and mining industry employment in TZ destination				0.108 (1.76)		
Manufacturing industry employment in TZ destination	0.191 (3.61)			0.118 (2.09)	0.146 (3.55)	0.254 (9.21)
Construction industry employment in TZ destination		0.158 (1.77)	0.433 (4.74)	0.166 (1.73)		
Wholesale and retail trade industry employment in TZ destination			0.248 (3.65)			
Accommodation and food services industry employment in TZ destination					0.238 (4.87)	0.104 (3.06)
Transport and warehousing industry employment in TZ destination	0.199 (2.99)					
Financial and insurance services industry employment in TZ destination		0.110 (1.95)				
Media, professional, scientific and ICT services industry employment in TZ destination			-0.255 (3.06)			0.093 (1.91)
Public administration industry employment in TZ destination				0.122 (2.92)		
Real estate, administration and other services industry employment in TZ destination		0.257 (2.34)				
Number of estimated parameters	***5***	***5***	***4***	***5***	***5***	***5***
Sample size	***200***	***200***	***200***	***209***	***209***	***555***
Restricted log likelihood	***-962.44***	***-962.44***	***-962.44***	***-1,005.75***	***-1,005.75***	***-2,670.76***
Final log likelihood	***-665.37***	***-681.93***	***-651.11***	***-679.11***	***-686.50***	***-1,819.25***
AIC/n	***6.70***	***6.87***	***6.55***	***6.55***	***6.62***	***6.57***

The SA3 model results show that the generalised cost (GenCost), the generalised cost to LCV driver home/depot (GenCostHH), and the logsum for the lower level TZ model are statistically significant in all models – taking into account that the GenCostHH could not be included in the stop 1 or 2 models since it was highly correlated to the GenCost. In branch 5 and branch 4, the results suggest that the sensitivity towards the GenCostHH increase as the LCV driver reaches its last stop, i.e., the LCV drivers try to get closer to their household (origin) towards the end of their tour-based trip. The CBD area dummy variable was statistically significant in several models. In branch 5 it was significant and positive in stop 1 and 2, suggesting drivers are more likely to stop in the CBD at the beginning of their tour. The same can be inferred from branch 4 results, where the CBD is positive in stop 1 and 2 and negative in stop 4. Similar results were found in branch 2, with a positive CBD parameter in stop 1 and negative in stop 2.

Table 6: SA3 Model Results for Branch 5 and 4 – mean (t value)

Attribute	Branch 5					Branch 4			
	Stop 1	Stop 2	Stop 3	Stop 4	Stop 5	Stop 1	Stop 2	Stop 3	Stop 4
Ln(Generalised cost from previous TZ[2] stop to alternative SA3 of the current stop)	-5.184 (4.92)	-2.461 (23.26)	-2.123 (17.39)	-2.407 (17.83)	-2.400 (33.20)	-2.284 (14.81)	-2.082 (15.44)	-1.903 (11.52)	-1.623 (8.25)
Ln(Generalised cost to LCV driver home/depot)			-0.876 (5.83)	-1.069 (6.90)	-1.899 (20.77)			-1.133 (5.77)	-2.140 (10.41)
If the destination SA3 is in the CBD area (1,0)	0.456 (2.32)	0.557 (2.78)				0.535 (2.03)	0.218 (0.83)		-0.497 (1.42)
Industry type manufacturing (1,0)				0.887 (3.08)					
Industry type transport and warehousing (1,0)		0.509 (1.72)	-0.878 (1.87)						
Industry type financial and insurance services (1,0)		1.329 (1.90)			1.177 (3.01)				
Industry type education and training (1,0)					0.887 (1.87)				
Female driver (1,0)				-0.901 (1.77)					
Ln(Household income of LCV driver in '000$AUD)	0.533 (2.39)								
Logsum of bottom TZ model	1.000 (6.56)	0.636 (4.49)	0.787 (7.22)	0.526 (4.90)	0.778 (11.02)	0.862 (3.90)	0.937 (3.93)	0.348 (1.95)	0.389 (1.47)
Number of estimated parameters	***4***	***5***	***4***	***5***	***5***	***3***	***3***	***3***	***4***
Sample size	***232***	***232***	***232***	***232***	***724***	***106***	***106***	***106***	***106***
Restricted log likelihood	***-925.44***	***-912.18***	***-916.69***	***-925.44***	***-2,927.17***	***-385.58***	***-391.02***	***-388.34***	***-382.76***
Final log likelihood	***-463.51***	***-498.25***	***-421.13***	***-405.49***	***-1,081.66***	***-220.07***	***-225.32***	***-187.57***	***-161.11***
AIC/n	***4.03***	***4.34***	***3.66***	***3.54***	***3.00***	***4.21***	***4.31***	***3.60***	***3.12***

Table 7: SA3 Model Results for Branch 3, 2 and 1 – mean (t value)

Attribute	Branch 3			Branch 2		Branch 1
	Stop 1	Stop 2	Stop 3	Stop 1	Stop 2	Stop 1
Ln(Generalised cost from previous TZ[2] stop to alternative SA3 of the current stop)	-4.420 (4.47)	-4.327 (4.42)	-1.612 (13.40)	-4.812 (4.83)	-2.103 (21.87)	-2.251 (30.14)
Ln(Generalised cost to household to LCV driver home/depot)			-1.863 (13.94)			
If the destination SA3 is in the CBD area (1,0)				0.354 (1.81)	-0.665 (2.60)	
Industry type manufacturing (1,0)						
Industry type construction (1,0)				0.422 (1.94)		0.369 (3.12)
Industry type education and training (1,0)					-1.440 (1.73)	
Industry type health care and social assistance (1,0)					-1.285 (1.55)	
Ln(Household income of LCV driver in '000$AUD)	0.485 (2.30)	0.474 (2.27)		0.499 (2.35)		
Logsum of bottom TZ model	0.753 (4.97)	1.000 (7.44)	0.669 (3.85)	0.539 (3.76)	0.794 (5.18)	0.953 (11.24)
Number of estimated parameters	***3***	***3***	***3***	***5***	***5***	***3***
Sample size	***200***	***200***	***200***	***209***	***209***	***555***
Restricted log likelihood	***-797.80***	***-794.06***	***-790.25***	***-821.75***	***-837.53***	***-2,234.07***
Final log likelihood	***-502.84***	***-469.61***	***-380.91***	***-482.13***	***-499.20***	***-1,750.97***
AIC/n	***5.06***	***4.73***	***3.84***	***4.66***	***4.82***	***6.32***

[2] For the first SA3 stop, the generalised cost is calculated from the TZ of the LCV driver's home/depot. For Stop 2 to Stop 5, the generalised cost is calculated from the TZ of the immediate previous stop.

The number of stops model results are presented in Table 8. The results show that LCVs are more likely to do one stop tour-based trips on weekends relative to 2 or more stops. The age of drivers seems to have a significant and positive influence on branch 2, which suggests that older drivers are more likely to do 2-stop tours, relative to the other number of stops. The logsums in the number of stops model, which represent the lower-level models were significant for all branches. The results show that the influence of the bottom branch has a higher influence on the 4-stop branch, followed by the 5-stop branch, 3-stop branch. These results suggest that there is a relationship between the number of stops and the destination choices, which is more relevant when the number of stops is higher.

Table 8: Tour type model describing the number of stops per tour

Variable	Mean (t value)
Alternative specific constant branch 2	-2.711 (3.20)
Alternative specific constant branch 3	0.019 (0.05)
Alternative specific constant branch 4	-0.149 (0.43)
Alternative specific constant branch 5	0.771 (2.28)
Age of LCV driver branch 2	0.528 (2.33)
Industry type education and training branch 1 (1,0)	1.293 (3.00)
Industry type accommodation and food services branch 1 (1,0)	0.503 (2.93)
Industry type financial and insurance services branch 5 (1,0)	-0.728 (4.29)
Logsum branch 1 stop	0.006 (1.69)
Logsum branch 2 stops	0.003 (1.84)
Logsum branch 3 stops	0.145 (2.34)
Logsum branch 4 stops	0.518 (3.93)
Logsum branch 5 stops	0.190 (3.59)
Weekend branch 1 (1,0)	0.623 (3.24)
Number of estimated parameters	**14**
Sample size	**1302**
Restricted log likelihood	**-1963.77**
Final log likelihood	**-1859.42**
AIC/n	**2.88**

5.1 Generalised cost elasticities

In this research, we are interested in obtaining a better understanding of how sensitivity towards the generalised cost varies across branches and throughout the tour-based trip. The parameter estimates presented above cannot be compared directly across models, so the elasticities for the normalised generalised costs are presented in Table 9 and Table 10 for the TZ models; and the elasticities for the generalised costs in Table 11 and Table 12 for the SA3 models.

The TZ models use the normalised generalised cost, which has a value close to 0 as it is calculated relative to the centre of the SA3. In branch 5, the sensitivity towards the normalised generalised cost increase as the trip progresses – although it is slightly higher in stop 1 than in stop 2. That is, the results show that there is 6.3% decrease in the probability of choosing a TZ in the first stop if the normalised generalised cost is 10% higher; 6.1% in the second stop; 7.0% in the third stop; 7.6% in the fourth stop; and 8.6% in the last stop. The interaction terms show that for the first stop, if the driver is male the sensitivity is a bit lower, namely, there is a 5.5% decrease in the probability of choosing a TZ when the normalised generalised cost is 10% higher. If the driver is female, then the sensitivity is higher, with a decrease of 10.6% in the probability when the cost increases by 10%. The interaction with gender is also significant in stop 5 and in branch 3 stop 1; always suggesting a higher sensitivity by female drivers towards the normalised generalised cost. It is interesting to note that the average sensitivity in each branch seems to decrease with the number of stops, i.e., the elasticity in branch 5 varies between -

0.61 and -0.86, while in branch 1 it is equal to -0.47. It seems that LCVs are more sensitive towards the normalised generalised cost when choosing a TZ if their tour-based trip is longer (with more stops).

Table 9: TZ Model Elasticities for Branch 5 and 4

	Branch 5					Branch 4			
	Stop 1	Stop 2	Stop 3	Stop 4	Stop 5	Stop 1	Stop 2	Stop 3	Stop 4
Normalised generalised cost	-0.63	-0.61	-0.70	-0.76	-0.86	-0.67	-0.58	-0.60	-0.75
Normalised generalised cost if industry type manufacturing				-0.85					
Normalised generalised cost if industry type construction					-0.91				
Normalised generalised cost if industry type accommodation and food services					-0.86				
Normalised generalised cost if industry type financial and insurance services			-0.63						
Normalised generalised cost if gender male	-0.55				-0.78				
Normalised generalised cost if gender female	-1.06				-1.33				

Table 10: TZ Model Elasticities for Branch 1, 2 and 3

	Branch 3			Branch 2		Branch 1
	Stop 1	Stop 2	Stop 3	Stop 1	Stop 2	Stop 1
Normalised generalised cost	-0.45	-0.49	-0.74	-0.55	-0.50	-0.47
Normalised generalised cost if industry type agriculture and mining						-0.44
Normalised generalised cost if industry type accommodation and food services	-0.47					
Normalised generalised cost if industry type health care and social assistance					-0.42	
Normalised generalised cost if gender male	-0.36					
Normalised generalised cost if gender female	-1.08					

The elasticity results for the SA3 models are calculated relative to the generalised cost. The results for branch 5 suggest that during the tour-based trip LCVs become less sensitive towards the generalised cost. Namely, if the generalised cost increase by 10%, there is a 18.8% reduction in the probability of choosing a SA3 in stop 1, 17.0% in stop 2, 14.3% in stop 3, 14.4% in stop 4, and 13.6% in stop 3. The drop in the elasticity in stop 3 can also be attributed to the generalised cost to the LCV driver home/depot, which was significant in this model and gained importance as the trip progressed. That is, in stop 3 the probability of choosing an SA3 decreases by 5.7% when the generalised cost to the LCV driver home/depot increases by 10%; decreases by 6.6% in stop 4; and by 10.0% in stop 5.

The results for branch 4 and branch 3 also suggest that as the trip progresses, the sensitivity towards the generalised cost decreases and the generalised cost to the LCV driver home/depot gains importance. In branch 2 and branch 1, the generalised cost to the LCV driver home/depot was not included as it was highly correlated with the generalised cost, so the results are a bit different. In branch 2, the sensitivity towards the generalised cost is actually higher in stop 2 than in stop 1: if there is an increase in the generalised cost by 10%, the probability to choose a SA3 decreases by 13.9% in stop 1 and 16.5% in stop 2. Gender was only significant in the branch 5 stop 4 model, suggesting that female drivers are more sensitive towards the generalised cost than male drivers.

Table 11: SA3 Model Elasticities for Branch 5 and 4

	Branch 5					Branch 4			
	Stop 1	**Stop 2**	**Stop 3**	**Stop 4**	**Stop 5**	**Stop 1**	**Stop 2**	**Stop 3**	**Stop 4**
Generalised cost	-1.88	-1.70	-1.43	-1.44	-1.36	-1.72	-1.51	-1.18	-0.97
Generalised cost to LCV driver home/depot			-0.57	-0.66	-1.09			-0.70	-1.28
Generalised cost if industry type manufacturing				-1.54					
Generalised cost if industry type transport and warehousing		-1.74	-1.38						
Generalised cost if industry type financial and insurance services		-1.72			-1.37				
Generalised cost if industry type education and training					-1.37				
Generalised cost if gender male				-1.39					
Generalised cost if gender female				-1.95					

Table 12: SA3 Model Elasticities for Branch 1, 2 and 3

	Branch 3			Branch 2		Branch 1
	Stop 1	**Stop 2**	**Stop 3**	**Stop 1**	**Stop 2**	**Stop 1**
Generalised cost	-1.77	-1.63	-1.12	-1.39	-1.65	-1.71
Generalised cost to LCV driver home/depot			-1.29			
Generalised cost if industry type construction				-1.47		-1.82
Generalised cost if industry type education and training					-1.62	
Generalised cost if industry type health care and social assistance					-1.63	

6 Generalised cost simulation

Using the model results for each branch and each stop, we simulated the probability to choose a TZ and a SA3 given different normalised generalised costs (nGenCost) and generalised costs (GenCost), respectively. We simulated scenarios where drivers had to choose between 4 different TZ or SA3s; that is why the probabilities to choose the TZ or SA3 will be equal to 25% when the generalised cost and normalised generalised cost are equal to the sample average (refer to Table 3). The simulated results for the TZ choice are presented in Figure 8. These graphs show that for branch 2 the sensitivity towards nGenCost is significantly higher in stop 1 than stop 2, particularly for negative nGenCosts (i.e., when the TZ is closer to the origin than the chosen SA3 centre). However, for branch 3 and 5 the opposite is true, where the sensitivity towards nGenCost is lower in stop 1 and increases in each stop – reaching the maximum sensitivity in the last stop. In branch 4 the first three stops have relatively similar sensitivities towards nGenCost, which increases in the last stop. In conclusion, it seems that for tour-based trips with at least three stops, LCVs are more sensitive towards the time and distance to the TZ towards the end of their trip, given they already chose the SA3.

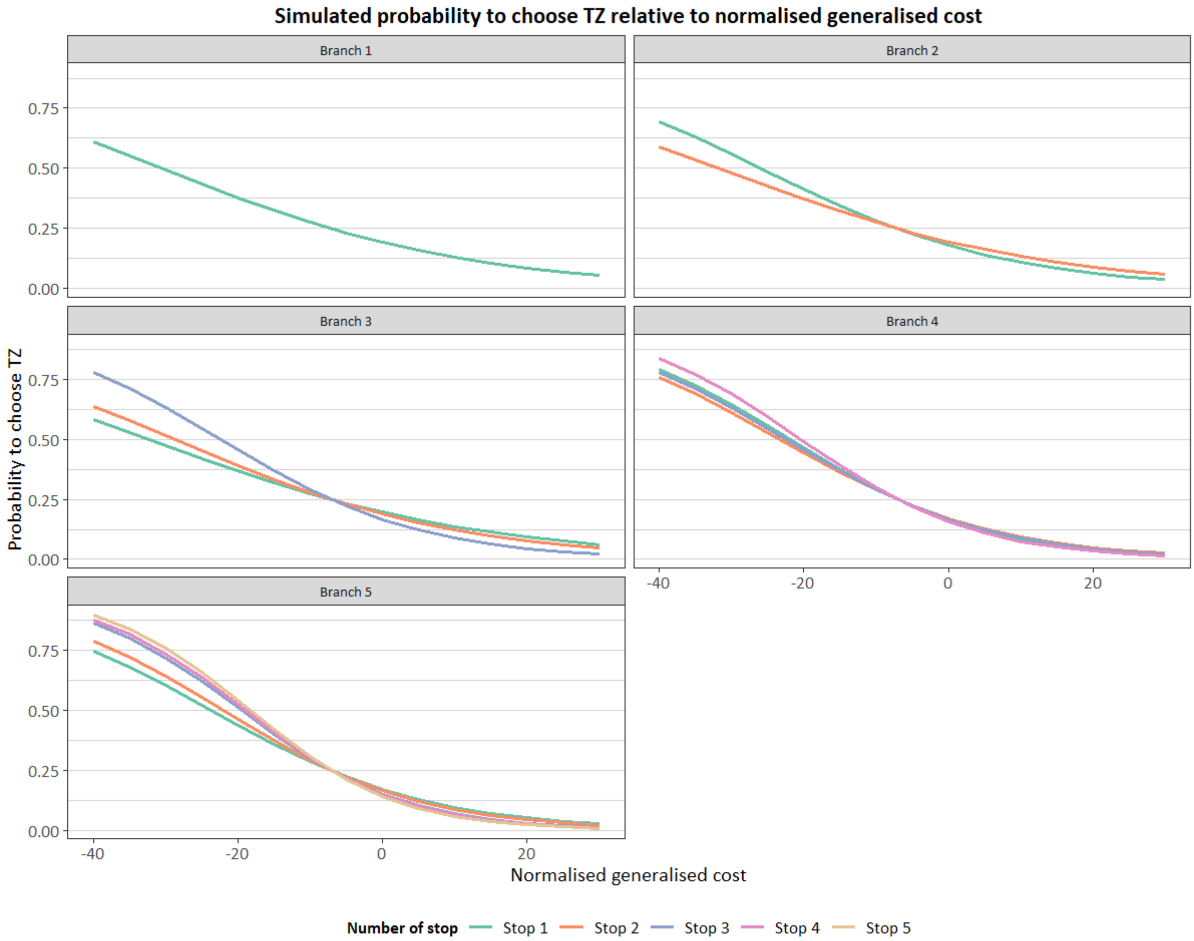


Figure 8: Simulated probability to choose TZ relative to normalised generalised cost

The simulated results which graphically represent the influence of the generalised cost in the SA3 choice are presented in Figure 9. The results for all branches show a higher sensitivity to the generalised cost for the first stop, which decreases in each stop. This is an important finding that should be complemented with the fact the time and distance to the LCV driver home/depot gains importance in the last three stops. These results are suggesting that LCVs are more sensitive towards the time and cost to the next SA3 in the first stops and decreases as they complete their trip. However, when they are far enough from their home/depot - which usually occurs in stop 3[3] - the time and distance back to their home/depot starts weighting in their decision-making.

[3] After the third stop the generalised cost between the origin and the alternative SA3, and the generalised costs between the origin and the household are no longer correlated.

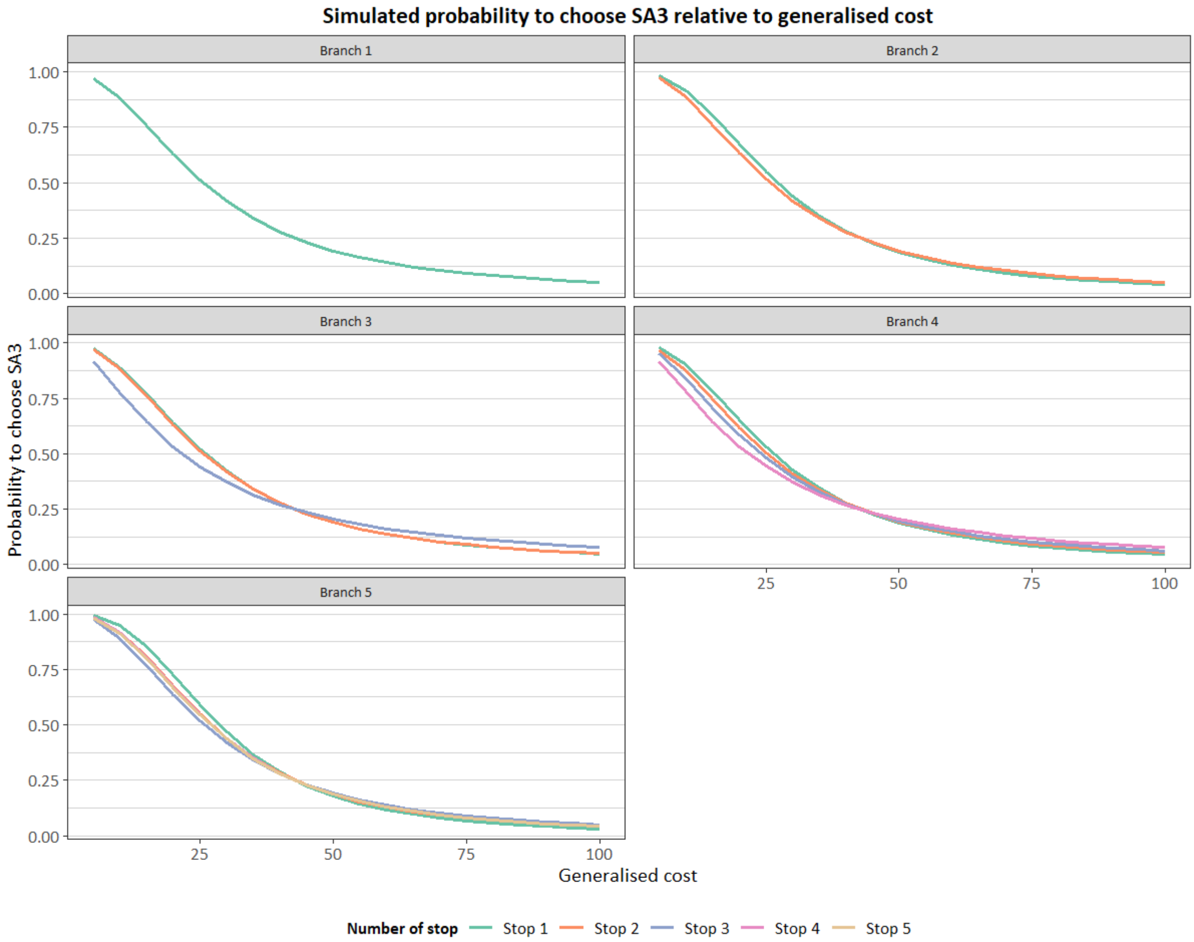


Figure 9: Simulated probability to choose SA3 relative to generalised cost

7 Discussion and conclusions

The current literature on tour-based trips of light commercial vehicles is somewhat limited. Usually, destination choices made by LCV drivers pool the data from different stops and different tour types (defined by the total number of stops). The objective of this research was to get a better understanding of preferences in different tour types and in different number of stops within the tour. The data used contained information of the LCV industry and household of the driver, driver characteristics (age, gender, income), and destination characteristics such as time and distance from the current location, employment and populations density. Since time and distance are highly correlated but they do contribute different measures (such as traffic or number of highways), they were combined into a generalised cost.

The choices studied in this work had three different levels: the first one is the decision of the number of stops, which defines the branches or the tour-based trip type; the second one the SA3 destination choice (which is a relatively large spatial unit defined in Sydney); and the third one the TZ destination choice (a smaller spatial unit) given the SA3 was already chosen. A separate TZ and SA3 destination choice model was estimated for each branch and each number of stop. The elasticities were calculated for the generalised costs and simulations were presented to show the changes in the probability to choose a certain destination given variations in the generalised costs.

The results show that when choosing the SA3, LCV drivers tend to be more concerned about the generalised cost in the first stop, which decreases as the trip is completed reaching its lowest sensitivity to the generalised cost in the last stop. Oppositely, the importance of the generalised cost back to the driver's home/depot increased as the trip is completed reaching a maximum in the last stop. That is, it seems like LCV drivers do not mind going further away in the middle stops but do want to be closer to their home/depot as they reach the end of their trip. This is an important finding, which suggests that it is very important to separate destination decision-making based on the tour type and also the stage of the tour. This study found important differences across branches and across the number of stops in the same branch.

Future research should focus on updating these models using data post-COVID. Since the start of COVID in 2020, online purchases in Australia have seen an all-time high with a 57% increase, and shoppers expect their online shopping 28% higher than before the pandemic (Australia Post, 2021). In Sydney, most of these trips are done by light commercial vehicles so it can be expected that these models gain importance in the future in city planning and transport logistics. This research provides a guideline as to the importance of identifying the number of stop and the total number of stops in the tour-based trip when understanding destination choices.

8 References

ATAP. (2016). *Australian Transport Assessment and Planning Guidelines: PV2 Road Parameter Values*.

Australia Post. (2021). *Inside Australian Online Shopping - eCommerce Industry Report 2021*.

Batley, R. (2015). The Hensher equation: derivation, interpretation and implications for practical implementation. *Transportation*, *42*(2), 257–275. https://doi.org/10.1007/s11116-014-9536-3

Boarnet, M. G., Hong, A., & Santiago-Bartolomei, R. (2017). Urban spatial structure, employment subcenters, and freight travel. *Journal of Transport Geography*, *60*, 267–276. https://doi.org/10.1016/j.jtrangeo.2017.03.007

Comi, A., Nuzzolo, A., & Polimeni, A. (2021). Aggregate delivery tour modeling through AVM data: experimental evidence for light goods vehicles. *Transportation Letters*, *13*(3), 201–208. https://doi.org/10.1080/19427867.2020.1868178

Crainic, T. G., Ricciardi, N., & Storchi, G. (2009). Models for Evaluating and Planning City Logistics Systems. *Transportation Science*, *43*(4), 432–454. https://doi.org/10.1287/trsc.1090.0279

Ellison, R. B., Teye, C., & Hensher, D. A. (2017). Modelling Sydney's light commercial service vehicles. *Transportation Research Part A: Policy and Practice*, *96*, 79–89. https://doi.org/10.1016/j.tra.2016.12.002

Hensher, D. A., Ho, C. Q., Liu, W., Wei, E., Ellison, R., Schroeckenthaler, K., Cutler, D., & Weisbrod, G. (2020). MetroScan: A quick scan appraisal capability to identify value adding sustainable transport initiatives. *Sustainability Special Issue: Transport Systems for Sustainability: Policy, Planning and Exploitation*, *12*(19). https://doi.org/10.3390/SU12197861

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). *Applied Choice Analysis - Second Edition*. Cambridge University Press. https://doi.org/10.1017/CBO9780511610356

Ho, C., & Mulley, C. (2013). Tour-based mode choice of joint household travel patterns on weekend and weekday. *Transportation*, *40*(4), 789–811. https://doi.org/10.1007/s11116-013-9479-0

Holguín-Veras, J., Thorson, E., Wang, Q., Xu, N., González-Calderón, C., Sánchez-Díaz, I., & Mitchell, J. (2013). Urban Freight Tour Models: State of the Art and Practice. In *Freight Transport Modelling* (pp. 335–351). Emerald Group Publishing Limited. https://doi.org/10.1108/9781781902868-017

Hunt, J. D., & Stefan, K. J. (2007). Tour-based microsimulation of urban commercial movements. *Transportation Research Part B: Methodological*, *41*(9), 981–1013. https://doi.org/10.1016/j.trb.2007.04.009

Hutchinson, B. G. (1974). Estimating urban goods movement demands. *Transportation Research Record: Journal of the Transportation Research Board*, *496*, 1–15.

Mackie, P. J., Wardman, M., Fowkes, A. S., Whelan, G., Nellthorp, J., & Bates, J. (2003). *Values of Travel Time Savings in the UK: Report to Department for Transport*.

Musolino, G., Polimeni, A., & Vitetta, A. (2018). Freight vehicle routing with reliable link travel times: a method based on network fundamental diagram. *Transportation Letters*, *10*(3), 159–171. https://doi.org/10.1080/19427867.2016.1241040

Nuzzolo, A., & Comi, A. (2013). Tactical and Operational City Logistics: Freight Vehicle Flow Modelling. *Freight Transport Modelling*, 433–451. https://doi.org/10.1108/9781781902868-021

Ogden, K. W. (1978). The distribution of truck trips and commodity flow in urban areas: A gravity model analysis. *Transportation Research*, *12*(2), 131–137. https://doi.org/10.1016/0041-1647(78)90052-7

Ruan, M., Lin, J. J., & Kawamura, K. (2012). Modeling urban commercial vehicle daily tour chaining. *Transportation Research Part E: Logistics and Transportation Review*, *48*(6), 1169–1184. https://doi.org/10.1016/j.tre.2012.06.003

Thoen, S., Tavasszy, L., de Bok, M., Correia, G., & van Duin, R. (2020). Descriptive modeling of freight tour formation: A shipment-based approach. *Transportation Research Part E: Logistics and Transportation Review*, *140*, 101989. https://doi.org/10.1016/j.tre.2020.101989

Wang, Q., & Holguín-Veras, J. (2008). Investigation of attributes determining trip chaining behavior in hybrid microsimulation urban freight models. *Transportation Research Record*, *2066*, 1–8. https://doi.org/10.3141/2066-01

Wisetjindawat, W., Sano, K., & Matsumoto, S. (2006). Commodity Distribution Model Incorporating Spatial Interactions for Urban Freight Movement. *Transportation Research Record: Journal of the Transportation Research Board*, *1966*(1), 41–50. https://doi.org/10.1177/0361198106196600106