International Journal of Transportation Science and Technology 7 (2018) 74–88

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

journal homepage: www.elsevier.com/locate/ijtst

Modeling route choice criteria from home to major streets: A discrete choice approach

Jose Osiris Vidana-Bencomo [a], Esmaeil Balal [b], Jason C. Anderson [c], Salvador Hernandez [d]

[a] Dept. of Civil Engineering, Universidad Autónoma de Ciudad Juárez, Av. Plutarco Elías Calles 1210, 32310 Ciudad Juárez, Chihuahua, Mexico
[b] Dept. of Civil Engineering, The University of Texas at El Paso, 500 West University Ave., El Paso, TX 79968, USA
[c] School of Civil and Construction Engineering, Oregon State University, 211 Kearney Hall, 1491 Southwest Campus Way, Corvallis, OR 97331, USA
[d] School of Civil & Construction Engineering, Oregon State University, 309 Kearney Hall, 1491 Southwest Campus Way, Corvallis, OR 97331, USA

ARTICLE INFO

Article history:
Received 22 May 2017
Received in revised form 17 November 2017
Accepted 19 December 2017
Available online 27 December 2017

Keywords:
Discrete choice model
Origin-destination matrix
Route choice
Traffic simulation
Transportation planning

ABSTRACT

A discrete choice model that consists of three sub-models was developed to investigates the route choice criteria of drivers who travel from their homes in the morning to the access point along the major streets that bound the Traffic Analysis Zones (TAZs). The first sub-model is a Nested Logit Model (NLM) that estimates the probability of a driver has or has no multiple routes, and if the driver has multiple routes, the route selection criteria are based on the access point's intersection control type or other factors. The second sub-model is a Mixed Logit (MXL) model. It estimates the probabilities of the type of intersection control preferred by a driver. The third sub-model is a NLM that estimates the probabilities of a driver selecting his/her route for its shortest travel time or to avoid pedestrian, and if the aim is to take the fastest route, the decision criteria is based on the shortest distance or minimum stops and turns. Data gathered in a questionnaire survey were used to estimate the sub-models. The attributes of the utility functions of the sub-models are the driver's demographic and trip characteristics. The model provides a means for transportation planners to distribute the total number of home-based trips generated within a TAZ to the access points along the major streets that bound the TAZ.

1. Backgound and motivation

This paper describes the development of a discrete choice model to predict the route choice criteria of drivers when they travel from homes to the access points of major streets, i.e., intersections along the main streets that bound the Traffic Analysis Zones (TAZs). This research is motivated by the network representation in a Microscopic Traffic Simulation (MTS) model as well as the need to convert an Origin-Destination (O-D) matrix from a Metropolitan Transportation Planning (MTP) model into the counterpart in a MTS.

In the past ten years, advances in MTS has enable several relatively new and commercial software to have the capability to simulate large urban networks faster than real-time. Along with this trend, several vendors have been marketing both MTP software and MTS software as compatible packages. Examples of such bundles are VISUM-VISSIM and

Peer review under responsibility of Tongji University and Tongji University Press.

https://doi.org/10.1016/j.ijtst.2017.12.002
2046-0430/

TransCAD-TransModeler. However, users of such tools must aware that MTP models are macroscopic while MTS models are microscopic. They represent objects in the network in different levels of details. Shelton et al. (2015) and Nava et al. (2012) have also developed tools that facilitate the so-called ''multiresolution simulation''. These tools are to convert a macroscopic MTP model into a mesoscopic traffic simulation model, and a mesoscopic traffic simulation model into a MTS model, respectively. The scope of this paper is about converting the O-D matrix of a MTP model directly into a MTS model.

MTP tools, being macroscopic in nature, model an urban area as multiple interconnected TAZs. Typically, each TAZ has one traffic generator (or centroid) that loads vehicles into the network and receives vehicles from the network. Users of MTS simulation usually demand a more detail network representation in order to simulate the traffic more realistically and to be able to identify local traffic problems. For example, a TAZ may be as small as 1 km $\times$ 1 km to 5 km $\times$ 5 km in area, depending on the population density, O-D trips and other factors. A TAZ in a MTP model is typically bounded by several arterials (or highways), has only one centroid and one connector that loads and receives vehicles between the centroid and one of the arterials. A MTS model usually codes this TAZ into several zones. Each zone its respective traffic generator that loads vehicles into and receives vehicles from an arterial. In this way, traffic load from this TAZ is split into several smaller zones, access and egress points along the major streets. This prevents an over-estimation of traffic demand and congestion at a single loading point, as often observed in MTP models.

The VISUM-VISSIM and TransCAD-TransModeler packages mentioned above, convert a TAZ in a MTP model into one traffic zone in the corresponding MTS model. To the best of the authors' knowledge, tool or algorithm has not been developed to facilitate the coding of a TAZ into multiple MTS zones, and to split the trips generated from and attracted to this TAZ into these zones. Users of MTS models always have to hard code and traffic zones, connectors and manually ''guestimate'' the traffic volume from/to each zone.

This paper assumes that, given the network topology of a TAZ, the major access points of traffic from the TAZ to the surrounding arterials can easily be identified. The objective of this research was to develop a disaggregated discrete choice model (DCM) to analyze drivers' route choice preferences when they select the routes from their homes to the access points of a city's transportation network. To make the model transferable, we have designed the choice set to be the route choice criteria (instead of specific routes in a TAZ). The model's attributes are based on socioeconomic characteristics and morning trip habits of the drivers. The data used to develop the DCM was obtained from a stated preference survey described in this paper.

The outline of this paper is as follows. After explaining the background and motivation, the next section of this paper reviews the application of DCMs in route choice. The subsequent section covers the model's framework and the fundamental concepts of DCM. This is followed by the description of the data used, and the model development process. The developed model was than applied to a TAZ in a case study, before making the conclusion.

2. Literature review

This section reviews the application of DCM and its variants in route choice analysis. DCMs have traditionally been used to mimic real life decisions, and in the transportation domain travel demand estimation, especially in mode choice. Wan et al. (2009) calibrated a Multinomial Logit (MNL) model to represent a commuter's route choice as a function of land use and transportation infrastructure. Yang et al. (2013) used cross-NLM and traditional NLM to describe the joint choice of residential location, travel mode, and departure time as functions of housing cost, travel time, travel cost and socio-demographic characteristics of the traveler. Song et al. (2015) analyzed a traveler's modes choice among bus, metro, walking, cycling, taxi and driving automobile by constructing a NLM as functions of value of time, trip distance, among other factors. Brands et al. (2014) presented a NLM to predict a traveler's mode choice. The choice set included single modes (including driving car) and multiple combinations of public transportation modes. The generalized cost functions included travel time, trip distance, waiting and transfer time and mode specific constants.

DCMs have also been applied in public transportation. Mehta and Lou (2013) tested a MNL model, a NLM and a Mixed Logit (MXL) model to predict the seat belt use of school bus riders in Alabama. Three alternatives were considered in the choice set: wearing, not wearing, and improperly wearing seat belts. A student's choice probabilities of these alternatives are modeled as functions of the student's demographic characteristics, trip attributes, seat location and level of bus driver's involvement. The NLM and MXL model were recommended. Nassir et al. (2015) used a NLM model to estimate the traveler's choice of transit access (boarding) stop as functions of available transit modes (among bus, train and ferry), the available routes within each modes and the departure times.

In the context of this research, some of the most recent applications of DCM have focused on the route choice behavior under the influence of traffic information and network topology. Xu et al. (2010) developed a MNL model to estimate a traveler's route choice behavior under the provision of traveler information. The attributes included the traveler's socio-demographic characteristics, content of traffic information and departure time. Mai et al. (2015) proposed a route choice model in which the choice of path is modeled as a sequence of link choices. This nested recursive logit model used link travel time; turn penalties, as the attributes of the links.

3. Modeling framework

DCMs predict human decisions among a set of distinct alternatives. In this application, it is used to predict a driver's route choice criteria when selecting a route from his/her home to an access point along a major street at the boundary of the TAZ, in the morning peak hour. The choice set consists of factors affecting the driver's decision on his/her routes. The probability of a driver selecting a route choice criterion is modeled as a function of the availability of routes, his/her demographic characteristics, trip characteristics, and personal route choice preference (street and intersection characteristics).

It is assumed that all residents in the TAZ are well informed of the various routes, street characteristics and traffic conditions. Choices among the different alternative routes are based on his/her own implicit decision criteria.

3.1. Model structure

Three different sub-models were developed to capture the decision process of a driver. The reason three sub-models were estimated instead of one was the limitation of the sample size in the observed data. Fig. 1 illustrates the structure of the model. The node at the first level (level I) is a binary decision node where the driver decides (1) if he/she has no multiple routes to access the main streets (*No Multiple*) or, (2) he/she has multiple routes to access the main streets (*Multiple*). If the driver has multiple routes, he/she then makes a (level II) decision on the route based on (1) the type of intersection control (*Intersection*) or, (2) the factor that is not related to intersection control (*Not Intersection*). If the driver's choice is *Intersection*, the next (level III) decision is to select one out of three types of intersection control: *Traffic Signal*, *Four-Way Stop* or *Two-Way Stop* as his/her criterion. If the driver's choice is *Not Intersection*, he/she will select between two alternatives: *Fastest Route* or *Avoid Pedestrians*. Under the *Fastest Route* option (level IV), there are two more branches: *Shortest Distance* or *Minimum Stops and Turns*.

3.2. Fundamentals of discrete choice models

In the context of route choice, logit-based models have been successfully applied to several recent studies (Raveau et al., 2014; Mai et al., 2015; Tan et al., 2015; Dalumpines and Scott, 2017). Therefore, the current study adopts this discrete choice model for a route choice application. For this work, specifically, the attractiveness of an alternative is modeled based on a linear-in-parameters function (Washington et al., 2011):

$$U_{in} = \boldsymbol{\beta_i x_{in}} + \varepsilon_{in} \tag{1}$$

where U_{in} = utility of alternative i to driver n; $\boldsymbol{\beta_i}$ = row vector of estimated coefficients for alternative i; $\boldsymbol{X_{in}}$ = column vector of measured attribute values for alternative i for driver n; ε_{in} = disturbance term that attempts to capture the unobserved factors present in the data.

Taking Eq. (1), the probability of driver n choosing alternative i can be represented by the standard multinomial logit form (Song et al., 2015; Washington et al., 2011):

$$P_{in} = \frac{e^{U_{in}}}{\sum_{\boldsymbol{j}} e^{U_{jn}}} \tag{2}$$

There are different types of DCMs, with different assumptions. The MNL model is one of the most basic forms of DCMs. Its extensions, the NLM and MXL models were used in this research. The MNL model, NLM and MXL model are explained in the next three sub-sections.

3.3. Multinomial logit model

The MNL model assumes that each ε_{in} is independent and identically distributed and follows a Gumbel distribution. In a MNL model, the probability of user n choosing outcome i is Washington et al. (2011), McFadden (1981):

$$P_n(i) = \frac{e^{\beta_i \boldsymbol{X}_{in}}}{\sum_{\boldsymbol{I}} e^{\beta_I \boldsymbol{X}_{In}}} \tag{3}$$

With a set of observed data $\boldsymbol{X}_{in}$, the coefficients $\boldsymbol{\beta}_i$ may be estimated by the method of maximum likelihood. Two-tailed t-test is used to evaluate the significance of each of the estimated coefficients. The log-likelihood ratio test is employed to evaluate the significance of the log-likelihood value obtained during MNL model estimation. This test compares the log-likelihood value (LL) of the model with all parameters set to zero, excluding the constant, to the model with all estimated parameters. Often, these two models are called the restricted model (log-likelihood at zero, only the constant) and the unrestricted model (log-likelihood with all estimated parameter). The log-likelihood expression follows a χ^2 distribution and is represented as Washington et al. (2011):

$$\chi^2 = -2[LL(\beta_R) - LL(\beta_U)] \tag{4}$$

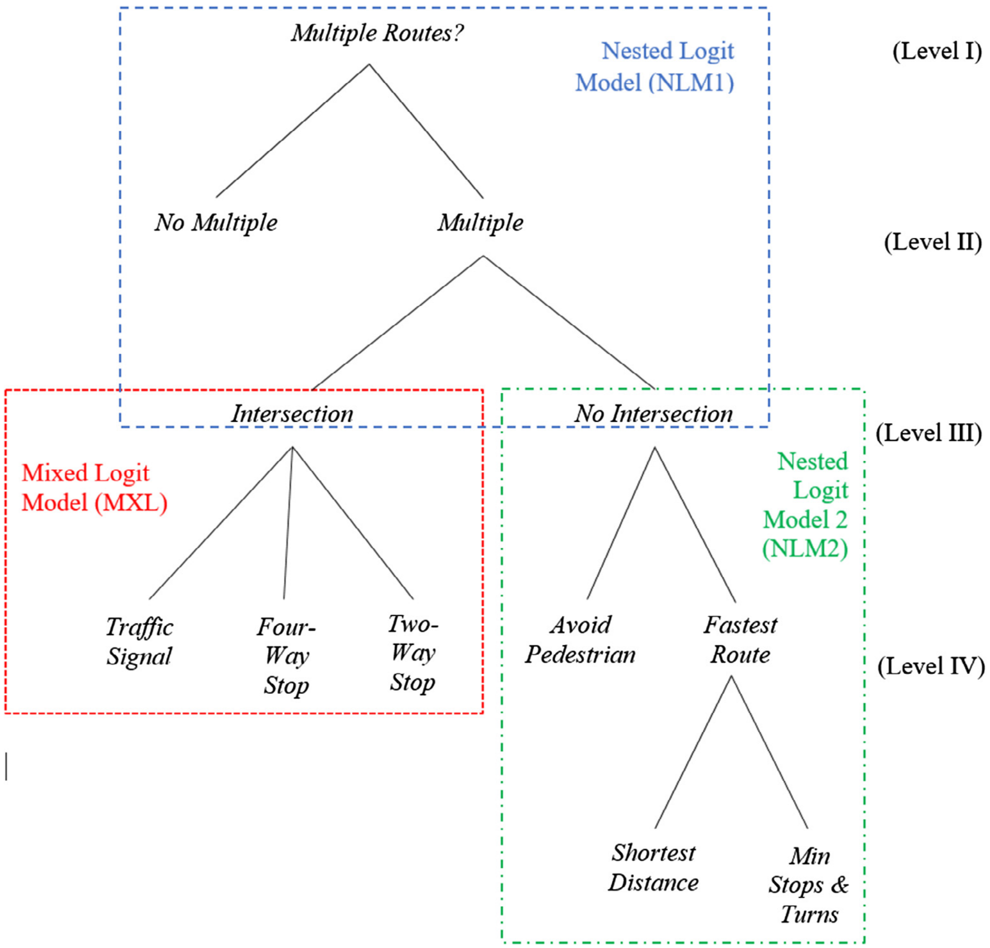

Fig. 1. Structure of discrete choice model.

Another common metric used to evaluate overall model fit is the McFadden Pseudo R-Squared value (also referred to as ρ^2):

$$\rho^2 = 1 - \frac{LL(\beta_U)}{LL(\beta_R)} \tag{5}$$

The McFadden Pseudo R-Squared value, ρ^2, is a scalar measure which varies between 0 and 1. However, unlike linear regression, values between 0.10 to 0.20 are said to have adequate fit, while values between 0.20 and 0.40 are said to have exceptional fit (McFadden, 1973, 1977).

Marginal effects are used to evaluate the impact of change in an attribute's value. The marginal effect represents the change in probability due to a one-unit increase in an attribute value. For a binary attribute, the marginal effect is computed as the difference in probability when an attribute (explanatory variable) changes from zero to one (Greene, 2012):

$$M_{\mathbf{X}_{inj}}^{P_n(i)} = \Pr[P_n(i) = 1|\bar{\mathbf{X}}_{(X_{inj})}, X_{inj} = 1] - \Pr[P_n(i) = 1|\bar{\mathbf{X}}_{(X_{inj})}, X_{inj} = 0] \tag{6}$$

where $\bar{\mathbf{X}}_{(X_{inj})}$ is the means of all other variables (i.e., all other variables are held constant) while X_{inj} changes from zero to one.

3.4. Nested logit model

The NLM groups the alternatives that are suspected to have shared error terms and/or unobserved values (ε_{in}) in a nested (sequential) decision tree. The probability expression is structured by unconditional and conditional cases that form the decision tree of the model. The probability of driver n selecting outcome i after making the decision on alternative j is computed as:

$$P_n(i) = \frac{e^{(\beta_i \boldsymbol{X}_{in} + \phi_i L_{in})}}{\sum_{\forall I} e^{(\beta_I \boldsymbol{X}_{In} + \phi_I LS_{In})}} \tag{7}$$

$$P_n(j|i) = \frac{e^{(\beta_{j|i} \boldsymbol{X}_n)}}{\sum_{\forall J} e^{(\beta_{J|i} \boldsymbol{X}_{Jn})}} \tag{8}$$

$$LS_{in} = \ln\left[\sum_{\forall J} e^{(\beta_{J|i} \boldsymbol{X}_{Jn})}\right] \tag{9}$$

where $P_n(i)$ = the unconditional probability of driver n selecting outcome i; $\boldsymbol{X}$ = vector of explanatory variables (i.e., measurable characteristics) used to determine the probability of outcome i; $\boldsymbol{\beta}$ = vector of estimated parameters based on $\boldsymbol{X}$; $P_n(j|i)$ = conditional probability of driver n selecting outcome j conditioned on the outcome being in outcome category i; LS_{in} = inclusive value (also referred to as the logsum); $\phi_{\boldsymbol{I}}$ = estimable parameter associated with $\boldsymbol{LS_{in}}$, in which $\mathbf{0} < \phi_i < 1$.

Lastly, to determine if the nested structure is the appropriate modeling framework, the inclusive value must be statistically different than 1. To assess this, a t-statistic is calculated to determine if the inclusive value, LS_{in}, is statistically different from 1 (Washington et al., 2011):

$$t^* = \frac{\beta_{LS_{in}} - 1}{SE(\beta_{LS_{in}})} \tag{10}$$

where t^* is a t-statistic based on the estimate of the inclusive value, $\beta_{LS_{in}}$, and the standard error of the inclusive value, $SE(\beta_{LS_{in}})$.

3.5. Mixed logit model

The MXL model, unlike the traditional MNL, can account for driver-specific variation due to variation within the data. This variation is often a result of not all data being available or a result of factors that are unseen to the analyst (commonly referred to as unobserved heterogeneity). To do this, $\boldsymbol{\beta}_i$ is given a mixing distribution (typically specified to be normal, but several distributions are tested for statistical significance), in which parameters are now permitted to vary across drivers. If this variation is not accounted for, model estimates and their corresponding inferences will be incorrect (Cerwick et al., 2014; Mannering et al., 2016; Anderson and Hernandez, 2017). Therefore, the choice probability based on the expression of the MNL model (Eq. (2)) can now be represented in terms of a mixing distribution (McFadden and Train, 2000; Washington et al., 2011; Mehta and Lou, 2013):

$$P_n(i|\theta) = \int_x \frac{e^{(\beta_i \boldsymbol{X}_{in})}}{\sum_{\forall I} e^{(\beta_I \boldsymbol{X}_{In})}} f(\boldsymbol{\beta}_i|\theta) d\beta_i \tag{11}$$

where $P_n(i|\theta)$ = the weighted outcome probability of outcome i conditional on $f(\boldsymbol{\beta}_i|\theta)$; $f(\boldsymbol{\beta}_i|\theta)$ = density function of $\boldsymbol{\beta}_i$ conditional on the distribution parameter θ, in which the distribution of $\boldsymbol{\beta}_i$ is specified by the analyst.[1]

4. Data

A questionnaire survey was conducted to collect data at the individual driver level for the estimation of the attribute coefficients in the utility functions. The survey instrument consisted of 16 questions in four sections to capture a respondent's demographic characteristics, morning trip characteristics, stated route choice preference and access point from home to the main street. Data from the first three sections were used to estimate the model coefficients while the answers in the fourth section were to check the consistency of earlier answers. The description of the variables and their categories are listed in Table 1.

The initial survey instruments were tested with 40 participants and revised before conducting the full survey from January to February 2011. The potential participants were approach by surveyors either in their residences, or at public areas (e.g., shopping malls). They were offered the options of fill out the forms themselves or be interviewed by surveyors.

A total of 1,133 El Paso residents participated in this survey, of which 484 participants (43%) said they decided on their routes based on the type of intersection control, 429 participants (38%) used criteria other than the type of intersection control, and finally, 218 participants (19%) answered that they had no alternate routes to travel to the main street. Among the subjects who choose *Intersection*, the percentage splits between the four types of intersection control are: 55% for signalized intersections, 20% for four-way stop control, 24% for two-way stop control and 1% roundabout and others. For those participants who selected their routes by criteria other than intersection control, the distribution of the criteria was: 55% based on shortest route, 20% based on minimum stops and turns, 19% based on avoiding school zone or pedestrians, and 5% depending

[1] The density function of $\boldsymbol{\beta}_i$ is what allows the parameters to vary across drivers; therefore, allowing for driver-specific variation. This distribution can be defined as any distribution (e.g., uniform, triangular, lognormal, etc.), but only a distribution with statistically significant standard deviations is used.

Table 1
Variables obtained from questionnaire survey.

Variable	Description and choices
Demographic	
Gender	Gender (1–male, 0–female)
Age	Age (1 < 25, 2–25 to 34, 3–35 to 49, 4–50 to 64, 5–65 and older)
Income	Annual household income (1–<$20 k, 2–$70 k to $89 k, 3–$20 k to $39 k, 4–$40 k to $69 k, 5– $90 k to $150 k, 6–>$150 k)
Education	Highest education level (1–high school, 2-community college, 3–undergraduate, 4–graduate)
Zip	Zip code (5 digits)
Household size	Household size (1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more)
No. of vehicles in household	Number of vehicles in the household (1–1, 2–2, 3–3, 4–4 or more)
Trip characteristics	
Purpose of weekday trip	Purpose of trips in a typical weekday morning (1–work, 2-school, 3-shopping/grocery, 4-gym, 5-other), may select more than 1 option
Stop in the neighborhood	Stop in the neighborhood before leaving the TAZ (1-at least a stop, 0-no stop)
Stop location	If Stop in the Neighborhood is "at least one stop", the type of location (1–gas station, 2-coffee shop, 3-daycare, 4-grocery store, 5-restaurant, 6-other), may select more than 1 option
Times leaving in the morning	Number of home-base trips out of TAZ in a typical weekday morning (1–1, 2–2, 3–3, 4–4 or more)
Times returning in the morning	Number of trips from outsize of TAZ to home in a typical weekday morning (1–1, 2–2, 3–3, 4–4 or more)
Route choice preference	
Route choice preference	Characteristics of route choice preference (1-intersection, 0-not intersection)
Intersection	If Preference is "intersection", the type of intersection control to access the main street (1-traffic signal, 2-four-way stop, 3-two-way-stop, 4-other)
Not intersection	If Preference is "not intersection", the other factors (1-shortest distance, 2-minimize stops, 3-avoid pedestrian, 4-drive through wider street, 5-other)
Access point	
Street names	Intersection to access the main street, denoted by the name of the neighborhood street and the name of the main street (text)
Ideal intersection control	Preferred type of intersection control in the neighborhood (1-traffic signal, 2-two-way stop, 3-four-way stop, 4-roundabout)

on other criteria. The demographic attributes of the participants are cross tabulated with their route choice preferences in Table 2. Values in parenthesis correspond to the total number of responses. The characteristics of the morning trips are cross tabulated with route choice preferences in the Table 3.

5. Discrete choice modeling

The data obtained from the survey were used to develop the different sub-models since in some cases of the MXL, the outcomes do not follow a nested structure; therefore, it would be inaccurate to fit a nested logit model. In some other cases there is a clear nested structure which requires that a nested logit model be fit which will be explained in next sections. The method of maximum likelihood was applied to estimate the coefficients of the utility functions, using LIMDEP (Greene, 2012). The criteria of *t*-statistics of between -1.28 and 1.28 (i.e., at 0.20 level of significance) was applied to select the attributes (Greene, 2012; Mai et al., 2015). The decision of 0.20 level of significance is due to the sample size. The selected attributes in the model and their descriptive statistics are listed in Table 4.

5.1. NLM1 – no multiple or multiple routes sub-model

The first (level I) decision node of the model (*Multiple* or *No Multiple*) was estimated as a sub-model. During the survey some drivers selected *No Multiple* because they may be using the same route every day without realizing the existence of alternate routes. Because of this, an Independence of Irrelevant Alternatives (IIA) violation could be present with the two subsequent options: *traffic control* and *non-traffic control*. To avoid this issue, a NLM was proposed as this sub-model to represent these choices. This NLM is denoted as NLM1. In this sub-model the first node contained the *No Multiple* and *Multiple* binary choices. Below the *Multiple* option, the choices of *traffic control* and *non-traffic control* were available. Table 5 lists all the independent variables with their estimated coefficients, t-statistics and marginal effects.

In order to justify the initial assumption of utilizing the NLM, Eq. (10) is used to calculate t^*. With an estimated value of 1.487 and standard error of 0.204, t^* is equal to 2.38; therefore, with over 95% confidence, the inclusive value is statistically significant from 1. This shows that the nested logit structure used in NLM1 is correct for this group of choices (see Washington et al., 2011) for further details on nested logit model fit criteria). In terms of significance of log-likelihood values, Eq. (4) results in a χ^2 of 61.49 with 18 degrees of freedom. This indicates with well over 99% confidence that the model with estimated parameters is of more significance than the model with only the constants.

The interpretation of the coefficients indicates that drivers with a household size of four are less likely to consider multiple routes and are also less likely to select the routes based on the type of intersection control. Marginal effects indicate that

Table 2
Results of questionnaire survey: demographic data.

Attribute	Category	Total response	Route choice preference			Intersection				Not Intersection				
			Intersection	Not intersection	No multiple	Traffic lights	4 Way stop	2 Way stop	Other	Shortest distance	Minimize stops & turns	Avoid pedestrians	Wider streets	Other
Gender	Male	44.0% (486)	18.6% (205)	16.1% (178)	9.3% (103)	50% (99)	20.2% (40)	27.3% (54)	2.55 (5)	59.9% (106)	19.8% (35)	14.1% (25)	0.0% (0)	6.2% (11)
	Female	56.0% (618)	23.8% (263)	22.1% (244)	10.1% (111)	59.8% (155)	18.5% (48)	20.8% (54)	0.8% (2)	52.1% (124)	20.6% (49)	22.3% (53)	0.8% (2)	4.2% (10)
Age	<25 yrs	17.2% (193)	7.1% (80)	7.0% (79)	3.0% (34)	53.8% (42)	20.5% (16)	24.4% (19)	1.3% (1)	61.5% (48)	17.9% (14)	17.9% (14)	1.3% (1)	1.3% (1)
	25 to 34 yrs	24.0% (270)	10.0% (112)	9.3% (104)	4.8% (54)	47.2% (51)	20.4% (22)	31.5% (34)	0.9% (1)	52.5% (53)	18.8% (19)	24.8% (25)	0.0% (0)	4% (4)
	35 to 49 yrs	26.0% (292)	10.4% (117)	10.9% (123)	4.5% (51)	54.3% (63)	19% (22)	25.9% (30)	0.9 (1)	52.5% (64)	19.7% (24)	18.9% (23)	0.0% (0)	9% (11)
	50 to 64 yrs	22.8% (257)	11.4% (128)	6.9% (77)	4.6% (52)	55.2% (63)	19.2% (22)	22.4% (28)	7% (3)	57.9% (44)	25.0% (19)	14.5% (11)	0.0% (0)	2.6% (2)
	65 & older	10.0% (113)	4.0% (45)	3.7% (42)	2.3% (26)	79.1% (34)	14.0% (6)	7.0% (3)	0% (0)	56.1% (23)	14.6% (6)	19.5% (8)	2.4% (1)	7.3% (3)
Income	< $20,000	15.3% (167)	6.5% (71)	5.7% (62)	3.1% (34)	55.4% (36)	20.0% (13)	23.1% (15)	1.5% (1)	59.7% (37)	21.0% (13)	17.7% (11)	0.0% (0)	1.6% (1)
	$20,000–$39,999	28.7% (313)	12.0% (131)	10.8% (118)	5.9% (640	56.9% (74)	21.5% (28)	20.8% (27)	0.8% (1)	46.1% (53)	21.7% (25)	29.6% (34)	0.9% (1)	1.7% (2)
	$40,000–$69,999	32.8% (358)	14.1% (154)	12.2% (133)	6.5% (71)	55.0% (83)	17.9% (27)	26.5% (40)	0.7% (1)	55.0% (72)	20.6% (27)	15.3% (20)	0.8% (1)	8.4% (11)
	$70,000–$89,000	13.8% (151)	6.0% (66)	5.4% (59)	2.4% (26)	49.2% (32)	20.0% (13)	30.8% (20)	0.0% (0)	54.2% (32)	22.0% (13)	16.9% (10)	0.0% (0)	6.8% (4)
	$90,000–$150,000	8.0% (87)	3.4% (37)	3.1% (34)	1.5% (16)	55.6% (20)	16.7% (6)	22.2% (8)	5.6% (2)	69.7 (23)	15.2% (5)	12.1% (4)	0.0% (0)	3.0% (1)
	over $150,000	1.5% (16)	0.5% (6)	0.5% (5)	0.5% (5)	66.7% (4)	0.0% (0)	16.7% (1)	16.7% (1)	100.0% (4)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
Education	High school	28.0% (311)	11.4% (126)	11.4% (127)	5.1% (57)	56.6% (69)	24.6% (30)	17.2% (21)	1.6% (2)	51.6% (64)	27.4% (34)	16.9% (21)	0.8% (1)	3.2% (4)
	Community Collegue	26.4% (293)	11.9% (132)	9.4% (104)	5.1% (57)	51.6% (66)	18.8% (24)	28.9% (37)	0.8% (1)	57.8% (59)	19.6% (20)	19.6% (20)	0.0% (0)	2.9% (3)
	Undergraduate	36.8% (409)	15.9% (176)	14.3% (159)	6.7% (74)	57.0% 998)	16.3% (28)	26.2% (45)	0.6% (1)	53.5% (85)	16.4% (26)	23.9% (38)	0.6% (1)	5.7% (9)
	Graduate	8.8% (98)	3.9% (43)	2.9% (32)	2.1% (23)	55.8% (24)	16.3% (7)	20.9% (9)	7.0% (3)	71.0% (22)	12.9% (4)	3.2% (1)	0.0% (0)	12.9% (4)

Household size	1	5.4% (61)	2.6% (290	1.7% (19)	1.2% (13)	70.4% (19)	18.5% (5)	11.1% (3)	0.0% (0)	55.6% (10)	22.2% (4)	11.1% (2)	0.0% (0)	11.1% (2)
	2	18.4% (206)	7.4% (83)	6.8% (76)	4.2% (47)	72.0% (59)	11.0% (9)	14.6% (12)	2.4% (2)	56.8% (42)	16.2% (12)	18.9% (14)	1.4% (1)	6.8% (6)
	3	19.5% (219)	9.1% (102)	7.1% (80)	3.3% (37)	57.6% (57)	17.2% (17)	23.2% (23)	2.0% (2)	60.8% (48)	19.0% (15)	13.9% (11)	1.3% (1)	5.15 (4)
	4	29.9% (336)	11.7% (131)	13.1% (147)	5.2% (58)	52.8% (67)	17.3% (22)	29.9% (38)	0% (0)	52.1% (75)	22.2% (32)	21.5% (31)	0.0% (0)	4.2% (6)
	5	19.7% (221)	9.2% (103)	7% (78)	3.6% (40)	42.2% (43)	28.4% (29)	28.4% (29)	1% (1)	57.7% (45)	19.2% (15)	20.5% (16)	0.0% (0)	2.6% (2)
	6	4.5% (50)	1.7% (19)	1.5% (17)	1.2% (14)	52.6% (10)	31.6% (6)	10.5% (2)	5.3% (1)	47.1% (8)	11.8% (2)	29.4% (5)	0.0% (0)	11.8% (2)
	7	1.3% (15)	0.4% (5)	0.4% (5)	0.4% (5)	40.0% (2)	20.0% (1)	20.0% (1)	20.0% (1)	60.0% (3)	40.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)
	8	0.9% (10)	0.3% (3)	0.3% (3)	0.4% (4)	66.7% (2)	0.0% (0)	33.3% (1)	0.0% (0)	66.7% (2)	0.0% (0)	33.3% (1)	0.0% (0)	0.0% (0)
	9	0.3% (3)	0.2% (1)	0.1% (1)	0% (0)	0.0% (0)	50.0% (1)	50.0% (1)	0.0% (0)	100.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
	10 or more	0.1% (1)	0.1% (0)	0.0% (0)	0% (0)	0.0% (0)	0.0% (0)	100% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
No. of vehicles in household	1	16.9% (190)	7.9% (89)	5.3% (60)	3.7% (41)	56.8% (50)	22.7% (20)	20.5% (18)	0.0% (0)	58.6% (34)	24.1% (14)	13.85 (8)	0.0% (0)	3.4% (2)
	2	50.0% (561)	22.2% (249)	18.8% (211)	9.0% (101)	56.0% (136)	16.9% (41)	25.5% (62)	1.6% (4)	54.5% (115)	17.5% (37)	22.3% (47)	0.9% (2)	4.7% (10)
	3	24.7% (277)	9.7% (109)	10.0% (112)	5.0% (56)	53.8% (57)	18.9% (20)	25.5% (27)	1.9% (2)	57.4% (62)	17.6% (19)	17.6% (19)	0.0% (0)	7.4% (8)
	4 or more	8.4% (94)	2.9% (32)	3.7% (42)	1.8% (20)	51.6% (16)	25.8% (8)	19.4% (6)	3.2% (1)	53.7% (22)	29.3% (12)	143.6% (6)	0.0% (0)	2.4% (%)

Number in brackets are frequencies of responses.

Table 3
Results of questionnaire survey: trip characteristics.

Attribute	Category	Response Total	Route choice preference			Intersection				Not Intersection				
			Intersection	Not intersection	No multiple	Traffic lights	4 Way stop	2 Way stop	Other	Shortest distance	Mini stops & turns	avoid pedestrians	Wider streets	Other
Purpose ot the Weekday Trip (more THAN one option)	Work	71.9% (813)	19.1% (365)	15.0% (287)	8.4% (161)	53.5% (192)	18.7% (67)	27.0% (97)	0.8% (3)	56.4% (159)	21.3% (60)	17.7% (50)	0.0% (0)	4.6% (13)
	School	45.5% (515)	12% (228)	10.0% (191)	5% (96)	48.9% (108)	18.6% (41)	32.1% (71)	0.5% (1)	51.0% (98)	21.9% (42)	20.3% (390	0.5% (1)	6.3% (12)
	Shopping / Grocery	31.5% (356)	8.4% (161)	6.6% (126)	3.6% (69)	55.3% (88)	16.4% (26)	25.8% (41)	2.5% (4)	50.4% (61)	21.5% (26)	22.3% (27)	0.8% (1)	5.0% (6)
	Gym	11.9% (135)	2.3% (44)	2.9% (55)	1.9% (36)	48.8% (21)	23.3% (7)	25.6% (11)	2.3% (1)	54.5% (30)	21.8% (12)	20.0% (11)	0.0% (0)	3.6% (2)
	Other	7.8% (88)	2.0% (38)	1.8% (34)	0.8% (16)	70.3% (26)	18.9% (7)	5.4% (2)	5.4% (2)	56.3% (18)	25.0% (8)	15.6% (5)	3.1% (1)	0.0% (0)
Stop in the neighborhood	No Stop	29.5% (334)	10.0% (113)	12.6% (143)	6.9% (78)	57.9% (66)	17.5% (20)	22.8% (26)	1.85 (2)	58.9% (83)	14.2% (20)	18.4% (26)	0.0% (0)	8.5% (12)
	Stop	70.5% (797)	32.8% (371)	25.3% (286)	12.4% (140)	54.5% (195)	19.6% (70)	24.6% (88)	1.4% (5)	53.7% (151)	22.8% (64)	19.9% (84)	0.7% 92)	3.2% (9)
Stop locations (more than one option)	Gas station	79.5% (627)	22.8% (301)	16.3% (215)	8.4% (111)	50.2% (146)	21.0% (61)	27.1% (79)	1.7% (5)	54.0% (114)	24.2% (51)	18.0% (38)	0.9% (2)	2.8% (6)
	Coffee shop	18.0% (142)	5.28% (69)	3.7% (49)	1.8% (24)	59.1% (39)	16.7% (11)	19.7% (13)	4.5% (3)	52.1% (25)	16.7% (8)	20.8% (10)	2.1% (1)	8.3% (4)
	Day care	9.5% (75)	2.9% (38)	2.1% (28)	0.7% (9)	56.4% (22)	20.5% (8)	20.5% (8)	2.6% (1)	53.8% (14)	19.2% (5)	23.1% (6)	0.0% (0)	3.8% (1)
	Grocery store	45.1% (356)	13.6% (180)	9.4% (124)	3.9% (52)	56.3% (99)	20.5% (36)	21.6% (38)	1.75 (3)	44.6% (54)	28.9% (35)	23.1% (28)	0.8% (1)	2.5% (3)
	Restaurant	11.5% (91)	4.2% (55)	17% (22)	1.1% (14)	43.4% (23)	26.4% (14)	24.5% (13)	5.7% (3)	45.5% (10)	18.2% (4)	22.7% (5)	0.0% (0)	13.6% (3)
	Other	3.9% (31)	1.0% (13)	0.9% (12)	0.5% (6)	53.8% (7)	30.8% (4)	0.0% (0)	15.4% (2)	33.3% (4)	50.0% (6)	8.3% (1)	0.0% (0)	8.3% (12)
Times leaving in the morning	0	4.8% (54)	2.0% (23)	1.8% (20)	1.0% (11)	56.6% (13)	26.1% (6)	13.0% (3)	4.3% (1)	61.1% (11)	11.1% (2)	5.6% (1)	5.6% (1)	16.7% (3)
	1	40.3% (456)	15.7% (178)	16.9% (191)	7.7% (87)	53.7% (94)	20.6% (36)	24.0% (42)	1.7% (3)	55.3% (104)	19.7% (37)	19.1% (36)	0.5% (1)	5.3% (10)
	2	32.7% (370)	15.7% (177)	11.2% (127)	5.8% (66)	49.7% (86)	19.7% (34)	30.1% (52)	0.6% (1)	58.1% (72)	18.5% (23)	18.5% (23)	0.0% (0)	4.8% (6)
	3	10.2% (115)	3.4% (39)	4.3% (49)	2.4% (27)	76.9% (30)	15.4% (6)	7.7% (3)	0.0% (0)	42.95 (21)	28.6% (14)	28.6% (14)	0.0% (0)	0.0% (0)
	4 or more	12% (136)	5.9% (67)	3.79% (42)	2.4% (27)	61.3% (38)	12.9% (8)	22.6% (14)	3.2% (2)	60.5% (26)	18.6% (8)	16.3% (7)	0.0% (0)	4.7% (2)
Times returning in the morning	0	38.3% (433)	14.7% (166)	17.0% (192)	6.6% (75)	56.6% (90)	19.1% (31)	24.1% (39)	1.2% (2)	57.4% (108)	19.1% (36)	16.5% (31)	1.1% (2)	5.9% (11)
	1	28.4% (321)	13.1% (148)	9.5% (108)	5.7% (65)	51.4% (76)	24.3% (36)	22.3% (33)	2.0% (3)	55.8% (58)	16.3% (17)	21.2% (22)	0.0% (0)	6.7% (7)
	2	19.5% (220)	8.8% (100)	6.7% (76)	3.9% (44)	56.3% (54)	12.5% (12)	31.3% (30)	0.0% (0)	51.9% (40)	26.0% (20)	13.0% (3)	0.0% (0)	1.3% (1)
	3	4.9% (55)	1.7% (19)	2.1% (24)	1.1% (12)	73.7% (14)	15.8% (3)	10.5% (2)	0.0% (0)	56.5% (13)	26.1% (6)	28.6% (14)	0.0% (0)	4.3% (1)
	4 or more	9.0% (102)	4.5% (51)	2.6% (29)	1.9% (22)	57.4% (27)	17.0% (8)	21.3% (10)	4.3% (2)	50.0% (15)	16.7% (5)	30.0% (9)	0.0% (0)	3.3% (1)

Number in brackets are frequencies of responses.

Table 4
Descriptive statistics of key variables in the model.

		Sub-Models					
		NLM1		MXL		NLM2	
Variable	Description (1 if yes, 0 otherwise)	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
GENDER	Male			0.415	0.493	0.417	0.493
AGE34	Age between 35 and 64	0.481	0.500				
AGE5	Age 65 or older			0.924	0.289		
HINC2	Annual household income between $70 K to $89 K					0.278	0.448
HINC6	Annual household income >$150 k	0.014	0.116				
HIGHSCH	High school education level	0.276	0.447	0.258	0.438		
GRAD	Graduate education level					0.067	0.251
HHSIZE4	Household size of four	0.299	0.458				
HHSIZE5	Household size of five			0.217	0.412		
HHSIZE6	Household size of six	0.042	0.202				
VEHICLE2	Two vehicles in the household					0.497	0.500
VEHICLE3	Three vehicles in the household					0.251	0.433
LEAVE3	Leaves home 3 times in the morning in a weekday	0.329	0.470				
LEAVE4	Leaves home 4 or more times in the morning in a weekday	0.104	0.306	0.838	0.277	0.123	0.329
RETURN1	Return home once in a weekday morning	0.282	0.450				
RETURN2	Return home twice in a weekday morning			0.311	0.463		
SCHOOL	Trip destination is school	0.455	0.498			0.447	0.497
DAYCARE	Regularly stops at daycare center	0.066	0.249				
COFFEE	Regularly stops at coffee shop			0.108	0.310	0.108	0.310
REST	Regularly stops at restaurant	0.077	0.267	0.107	0.309		
GASSTN	Regularly stops at gas station			0.615	0.486	0.510	0.500
GROCERY	Regularly stops at grocery store	0.317	0.466			0.291	0.454
OTHER	Regularly stops at other locations					0.027	0.164
INT1	Prefers traffic signal as the type of intersection control	0.480	0.500				
INT4	Prefers roundabouts as the type of intersection control	0.092	0.289			0.100	0.300

Table 5
Estimation results for NLM1 sub-model.

Variable	Description (1 if yes, 0 otherwise)	Coefficient	t-Stat	Marginal effects
No Multiple				
HINC6	Annual household income >$150 k	0.8481	1.505	0.131
HHSIZE4	Household size of four	−0.5033	−2.632	−0.078
LEAVE4	Leaves home 4 or more times in a morning	0.4512	1.372	0.070
REST	Regularly stops at restaurant	−0.6214	−1.813	−0.096
INT4	Prefers roundabouts	0.4700	1.753	0.073
Multiple				
HHSIZE6	Household size of six	−0.5974	−1.759	–
RETURN1	Returns home once in the morning	0.4221	2.543	–
GROCERY	Regularly stops at grocery store	0.5093	2.791	–
INT1	Prefers traffic signal	−0.3274	−1.996	–
Traffic control				
AGE34	Age between 35 and 64	0.2183	2.173	0.057
HHSIZE4	Household size of four	−0.4602	−3.143	−0.119
LEAVE3	Leaves home 3 times in a morning	0.3275	2.650	0.085
SCHOOL	Trip destination is school	0.1794	1.613	0.047
DAYCARE	Regularly stops at daycare center	0.3259	1.454	0.085
Non-traffic control				
HIGHSCH	High school education level	0.1887	1.507	0.046
LEAVE4	Leaves home 4 or more times in a morning	0.4135	1.782	0.101
REST	Regularly stops at restaurant	−0.9470	−3.389	−0.231
	Inclusive Value	*1.4870*	*7.27*	
	Number of observations	1103		
	Number of parameters	18		
	log-likelihood at zero	−1159.06		
	log-likelihood at convergence	−1128.32		
	χ^2	61.49		
	ρ^2	0.03		
	ρ^2 adjusted	0.02		

a driver with a household size of four has a 0.078 lower probability of selecting *No Multiple*. Similarly, a driver with a household of four has a 0.119 decrease in probability of selecting *Intersection*.

No Multiple and *non-traffic control* have the common attribute of leaving home four times or more in a morning. Marginal effects suggest that leaving home four times or more in a morning increases the probability of a driver selecting *No Multiple* by 0.070, while there is a 0.101 higher probability of selecting *Not Intersection.*

Drivers who regularly stop at a restaurant during his/her route are less likely to use *No Multiple* (0.096 lower probability, according to marginal effects) and *non-traffic control* (0.231 lower probability, based on marginal effects) as the route choice criteria. This may indicate that these drivers prefer to select their routes based on *Intersection*.

Drivers with annual household income higher than $150,000 are more likely to select *No Multiple* as their route choice (*ß* = 0.8481). Presumably their houses have limited access as a result of higher privacy. This attribute has one of the largest marginal effects, as marginal effects show a 0.131 higher probability of selecting *No Multiple* if the annual household income is greater than $150,000.

Drivers who prefer roundabouts in their neighborhood are more likely to select *No Multiple* (*ß* = 0.4700). This group shows interest in roundabout as the type of traffic control in the TAZ. For this variable, marginal effects show that preferring roundabouts results in a 0.073 increase in probability of a driver selecting *No Multiple.*

Drivers who are heading to schools in the morning are more likely to select access points to the main street based on the type of traffic control (*ß* = 0.1794). The model is probably capturing some safety concerns when drivers are dropping children off at schools. Marginal effects indicate that a trip to school will increase the probability of selecting *traffic control* choice by 0.047. This situation is similar to Drivers who regularly stop at day care centers (*ß* = 0.3259). In this case, a trip to a day care will increase the probability of selecting *traffic control* choice by 0.085.

Age is a contributing factor to the use of access points with traffic control. Drivers from 35 to 64 years old have preference by its use (*ß* = 0.2183). Marginal effects show that drivers from 35 to 64 years old have a 0.057 higher probability of selecting *traffic control.*

Drivers who leave home three times in a morning are more likely to select access points by the type of *traffic control* (ß = 0.3275). In this case, marginal effects suggest that drivers that leave home three times in a morning have a 0.085 higher probability of selecting *traffic control.*

Drivers with a high school education level are more likely to select their routes not by the type of *traffic control* (*ß* = 0.1887 for *non-traffic control*). This may suggest that drivers with only high school education are not as concern with safety issue.

5.2. MXL – traffic control type sub-model

The second sub-model is for the drivers who base their route choice decisions on the type of *traffic control*. The probability of a driver selecting among the three types of *traffic control* as the route choice criterion was modeled by a MXL model. The choice set consisted of traffic signal, two-way stop and four-way stop controls. The MXL model was selected because it addresses the weaknesses that can result in erroneous parameter estimates if underlying assumptions of MNL are not met (Washington et al., 2011) since in the case of the MXL, the outcomes do not follow a nested structure; therefore, it would be inaccurate to fit a nested logit model.

Table 6 shows the estimated attribute coefficients of the utility functions in the MXL sub-model, including the *t*-statistics and marginal effects. Applying Eq. (4), when comparing log-likelihood values of the fixed- and random-parameters models, results in a chi-square statistic of 5.66 with 2 degrees of freedom (the number of estimated random parameters). Based on a chi-square distribution, this suggests that the log-likelihood of the random parameters model is of more significance with well over 90% confidence.

Two variables were found to have heterogeneous effects across drivers and have estimated random parameters based on the statistical significance of the standard deviation: GENDER and GASSTN. The parameters of the probability density function $f(\boldsymbol{\beta}_i|\theta)$ of these variables are (-0.508, 5.079) for males and, 3.803) for drivers who regularly stops at gas stations before they leave the TAZ. Specifically, GENDER was found to have a normally distributed random parameter. With a mean of -0.508 and a standard deviation of 5.079, the normal distribution suggests that 46% of drivers have an estimated parameter mean greater than zero and 54% of drivers have an estimated parameter mean less than zero. In other words, 46% of male drivers are more likely to select the intersection with a traffic signal while 54% of male drivers are less likely. The randomness of this variable may be attributed to road network familiarity by males and unobservable factors related to congestion effects.

In the case of GASSTN, drivers who regularly stop at gas stations was found to have a normally distributed random parameter. With a mean of −0.485 and a standard deviation of 3.803, 45% of drivers who regularly stop at gas stations are more likely to select a route with a traffic signal and 55% are less likely. Usually gas stations are located at intersections controlled by a traffic signal. Again the variability of this variable may be attributed to unobservable factors related to lifestyle characteristics and congestion effect.

The coefficient of LEAVE4 indicates that drivers are more likely to use an intersection with traffic signal (*ß* = 1.831) if they leave home a minimum of four times in one morning. Further, drivers that leave home four or more times were also found to be less likely to select a route with a four-way stop intersection (*ß* = −1.084). Marginal effects show that the probability of using a signalized intersection will increase by 0.010 if a driver leaves home four or more times in a morning. Likewise, the

Table 6
Estimation results for MXL sub-model.

Variable	Description (1 if yes, 0 otherwise)	Coefficient	t-Stat	Marginal effects
Traffic Signal				
SIGNAL	Traffic signal control constant	0.934	3.58	–
GENDER	Male	−0.508* (5.079)**	−0.84 (1.55)	−0.017
AGE5	Age 65 or older	2.613	1.63	0.013
HHSIZE5	Household size of five	−1.184	−2.41	−0.027
LEAVE4	Leaves home 4 or more times in a morning	1.779	1.56	0.010
COFFEE	Driver who regularly stops at coffee shop	0.877	1.31	0.010
GASSTN	Regularly stops at gas station	−0.485* (3.803)**	−1.10 (1.91)	−0.029
Two-Way Stop				
TWSTOP	Two-way stop intersection control constant	−0.852	−3.97	–
REST	Regularly stops at restaurant	0.749	1.77	0.012
RETURN2	Returns home two times in the morning	0.679	2.29	0.030
Four-Way Stop				
AGE5	Age 65 or older	−1.053	−1.41	−0.004
HIGSCH	High school education level	−0.813	−2.52	−0.024
LEAVE4	Leave home four or more times in the morning	−1.092	−1.47	−0.006
	Number of observations	465		
	Number of parameters	15		
	log-likelihood at zero	−458.80		
	Log-likelihood of fixed parameters	−432.23		
	log-likelihood at convergence	−429.40		
	χ^2 (random parameters)	58.80 (5.66)		
	ρ^2	0.06		
	ρ^2 adjusted	0.05		

* Mean of random coefficient.
** Standard deviation of random coefficient.

probability of a driver selecting a four-way stop intersection will decrease by 0.006 if he/she leaves home four or more times in a morning.[2]

For the AGE5 variable, the statistics show that drivers 65 and older are more likely to go through an intersection with traffic signal (ß = 2.613). A possible explanation might be attributed to them being older and seeking safer intersections, which are generally controlled by signals.

Marginal effects suggest that If a driver is older than 65 years old, the probability that he/she uses a signalized intersection will increase 0.013 and decrease the probability of selecting a four-way stop intersection by 0.004. A possible reason is that these drivers are seniors who prefer safer intersections and are in no hurry to arrive to their destinations.

The bigger the household size, the less likely the driver will select a route that consists of a signalized intersection. This is especially true if the house hold size is five (ß = -1.184). This may be because the drivers need to drop off the rest of the household members, especially children, to different locations, and therefore they are less likely to use the access points that are controlled by signals in order to save time. For this factor, marginal effects indicate a 0.027 lower probability of selecting a traffic signal if the driver has a household size of five.

The more drivers buy coffee in the morning, the more likely they are to pass through a signalized intersection (ß = 0.877). These drivers probably stop at gas stations at corners of signalized intersections. Marginal effects suggest that an additional stop for coffee will contribute to an increase in the probability of using a signalized intersection by 0.010.

If drivers on his/her route need to stop at a restaurant, two-way stop intersection choice is most frequently used (ß = 0.749). This might be explained by neighborhood restaurants being located closer to this type of intersection. In terms of effect, marginal effects show that an additional restaurant stop will lead to an increase in the use of a four-way stop intersection by 0.012.

Drivers are more likely to go through a two-way stop intersection (ß = 0.679) if they decide to return home twice. This attribute may be capturing some driver behavior concerning minimize route impedance. A driver who switch from returning home other than two times to exactly two times in the morning will cause an increase in the probability of the use of a two-way stop control intersection by 0.030.

5.3. NLM2 – non-traffic control type sub-model

The NLM was used to model the data related to the choice of *non-traffic control.* The reason of using NLM was because there was an IIA violation between *Shortest Distance* and *Minimum Stops and Turns*. This estimated sub-model is denoted

[2] Being that this variable is present in more than one function, marginal effects on a specific outcome can be ambiguous and must be interpreted in such a manner (Anderson and Hernandez, 2017).

Table 7
Estimation results for NLM2.

Variable	Description of variable (1 if yes, 0 otherwise)	Coefficient	t-Stat	Marginal effects
Fastest Route				
GENDER	Male	0.582	0.27	–
COFFEE	Regularly stops at coffee shop	−0.273	−0.67	–
Shortest Distance				
SDIST	Shortest distance constant	1.196	4.20	–
VEHICLE3	Three vehicles in the household	0.687	1.88	0.105
LEAVE4	Leaves home 4 or more times in a morning	−0.539	−1.26	−0.082
SCHOOL	Trip destination is school	−0.348	−1.29	−0.053
OTHER2	Regularly stops at other locations	−1.846	−2.65	−0.281
Min Stops & Turns				
VEHICLE2	Two vehicles in the household	−0.518	−1.66	−0.075
GROCERY	Regularly stops at grocery store	0.884	3.01	0.128
Avoid Pedestrians				
AVOIDP	Avoid pedestrian constant	−0.948	−1.63	–
HINC2	Annual household income \$70 k to \$89 k	0.793	2.89	0.120
GRAD	Graduate education level	−1.750	−1.69	−0.264
GASSTN	Regularly stop at gas station	−0.384	−1.43	−0.058
INT4	Prefers roundabouts	−1.284	−2.06	−0.194
	Inclusive Value	*0.11080*	*0.29*	
	Number of observations	398		
	Number of parameters	14		
	log-likelihood at zero	−383.31		
	log-likelihood at convergence	−358.44		
	χ^2	49.74		
	ρ^2	0.06		
	ρ^2 adjusted	0.05		

as NLM2. In the NLM2 sub-model, the first decision was a binary choice between *Fastest Route* and *Avoid Pedestrian*. The more pedestrian crossing, the more stop time any driver has and that is why it is a binary choice between Fastest Route and Avoid Pedestrian since there is no possibility to get the Fastest Route with Pedestrians. If a driver selects *Fastest Route*, he/she has further options of *Shortest Distance* or *Minimum Stops and Turns* as the route choice criteria. Table 7 shows the estimated attribute coefficients of the utility functions in the NLM2 sub-model, including the t-statistics and marginal effects. In order to validate the initial assumption of the use of NLM, Eq. (10) is used to calculate t^*. With an estimated value of 0.1108 and a standard error of 0.3864, t^* is equal to -2.30. Therefore, will over 95% confidence, the nested logit structure is the correct form for this data. In terms of log-likelihood significance, with a chi-square statistic of 49.74 and 14 degrees of freedom, the log-likelihood of the model with estimated parameters is of more significance than the model with only constants with well over 99% confidence.

The interpretations of these attribute coefficients and marginal effects are discussed in the following paragraphs. A driver stopping at a coffee store is less likely to travel by a route that is the fastest ($ß = -0.273$). Males are more likely to drive to the access points that have the shortest travel time or fastest route ($ß = 0.582$). This behavior is probably due to the male's aggressiveness and impatience over female.

Drivers with three vehicles in the household showed to be more likely to choose the shortest distance ($ß = 0.687$). Marginal effects indicate that a one-unit increment of a household with three vehicles will increase the probability of drivers selecting the shortest distance by 0.105.

Drivers leaving home four or more times in the morning are less likely to choose the access points with the shortest distance ($ß = -0.539$). This situation could apply to drivers that make multiple trips and tend to have time to run errands. For this factor, marginal effects show that a one-unit increment in drivers that leave the home four times or more in the morning will decrease the probability of selecting the shortest distance by 0.082.

Drivers with final destination being school, or drivers that make other stops during his/her trip, are less likely to choose the access points with the shortest distance. The reason could be that because the detours to school and other destinations make them impossible to use the shortest distance paths to directly connecting their homes to the nearest major streets. With regard to impact, marginal effects suggest that drivers who take a detour to school or another location have a 0.053 and 0.281 lower probability of selecting the shortest distance path, respectively.

Drivers living in households with two vehicles are less likely to take the routes with minimum stops and turns ($ß = -0.518$). Specifically, a one-unit increase in drivers with two vehicles in the household will decrease the probability of them using this choice by 0.075.

Driver who regularly stop at grocery stores during morning trips are more likely to take routes which incur minimum stops and turns ($ß = 0.844$). A possible explanation is that drivers going grocery shopping usually like to avoid disturbing the contents in their vehicles' trunks. In terms of effect, marginal effects show that drivers stopping at a grocery store will increase the probability of drivers avoiding turns and stops by 0.128.

Drivers with annual household income from \$70,000 to \$89,000 are more likely to avoid streets with pedestrians ($ß$ = 0.793). Assuming that these drivers belong to middle class households, their houses are located away from parks and schools. Then, these drivers can easily avoid pedestrians during his/her home-based trips. Marginal effects indicate that a one-unit increase in middle class drivers (household) will increase the probability of traffic trying to avoid pedestrians by 0.120.

Drivers with a graduate education were less likely to choose a route that avoids pedestrians ($ß = -1.750$). A possible reason is that their decisions on the locations of home parcels already take this pedestrian friendly factor into consideration. This is the most impactful factors, as marginal effects show a 0.264 lower probability of selecting a route that avoids pedestrians.

Finally, commuters who stop to refuel gasoline are less likely to avoid pedestrians ($ß = -0.384$). The result is probably due to the fact that gas stations are located at intersections with the pedestrian presence.

6. Conclusions

This research has explored the use of three sub-models: one MXL model and two NLMs to quantify the contributions of a driver's demographic profile, morning trip characteristics on the probabilities of selecting intersection control types (signal, two-way stop, four-way stop), fastest route (shortest distance or minimum stops and turns), safer streets (avoiding pedestrian) as the criterion for route choice between his/her home to the access point along a major street that bound the TAZ.

Based on the three logit sub-models that have been fitted to the data collected in the City of El Paso, the following observations may be concluded concerning drivers' route choices from their homes to the major streets:

- Females, drivers older than 64 years old and drivers stopping at gas stations and coffee shops prefer to select their routes by the type of *traffic control*.
  - o Traffic signal control is preferred by drivers older than 64 years old and also by females.
  - o Two-way stop control is only preferred by drivers who stop at restaurants and those who return home two times in the morning.
  - o Four-way stop intersection is attractive to drivers who go to grocery stores in the morning. Drivers leaving home four times in the morning, with high school education, or older than 65 years old are less likely to use this type of intersection.
- Route with the shortest distance is preferred by males or drivers with three vehicles in the household.
- Drivers with graduate degrees or drivers stopping at gas stations tend not to avoid pedestrians. Middle income drivers are more likely take a route to avoid encountering pedestrians.

The estimated discrete choice model may be combined with the socioeconomic data of the TAZ (readily available in the metropolitan planning models) to predict the route choice preferences of drivers, from which the total trips generated by this TAZ during the morning commute may be distributed to multiple access point along the major streets at the boundary of this TAZ.

The research contributes to a better understanding of the route choice behavior between residential homes and the boundary of the TAZ, for home-based trips in the morning. The model developed is a step towards the conversion of a TAZ-based transportation planning O-D matrix into a higher resolution O-D matrix for used in microscopic traffic simulations.

References

Anderson, J., Hernandez, S., 2017. Roadway classifications and the accident injury severities of heavy-vehicle drivers. Anal. Methods Accid. Res. 15, 17–28.
Brands, T., de Romph, E., Veitch, T., Cook, J., 2014. Modelling public transport route choice, with multiple access and egress modes. Transp. Res. Procedia 1, 12–23.
Cerwick, D.M., Gkritza, K., Shaheed, M.S., Hans, Z., 2014. A comparison of the mixed logit and latent class methods for crash severity analysis. Anal. Methods Accid. Res. 3–4, 11–27.
Dalumpines, R., Scott, D.M., 2017. Determinants of route choice behavior: a comparison of shop versus work trips using the potential path area – gateway (PPAG) algorithm and path-size logit. J. Transp. Geogr. 59, 59–68.
Greene, W.H., 2012. Econometric Analysis. Pearson Education, New York, NY.
Mai, T., Fosgerau, M., Frejinger, E., 2015. A Nested Recursive Logit Model for Route Choice Analysis. Transp. Res. Part B: Methodological 75, 100–112.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.
McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior. In: Frontiers in Econometrics, pp. 105–142.
McFadden, D., 1977. Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments. Institute of Transportation Studies. University of California.
McFadden, D., 1981. Econometric models of probabilistic choice. In: Manksi, C.F., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge, MA, pp. 198–269.
McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. J. Appl. Econ. 15 (5), 447–470.
Mehta, G., Lou, Y., 2013. Modeling school bus seat belt usage: nested and mixed logit approaches. Accid. Anal. Prev. 51, 56–67.
Nassir, N., Hickman, M., Malekzadeh, A., Irannezhad, E., 2015. Modeling transit passenger choices of access stop. Transp. Res. Rec.: J. Transp. Res. Board 2493, 70–77.
Nava, E., Shelton, J., Chiu, Y., 2012. Analyzing impacts of dynamic reversible lane systems using a multiresolution modeling approach. In: Transportation Research Board 91st Annual Meeting.

Raveau, S., Guo, Z., Muñoz, J.C., Wilson, N.H.M., 2014. A Behavioral comparison of route choice on metro networks: time, transfers, crowding, topology and socio-demographics. Transp. Res. Part A: Policy Pract. 66, 185–195.

Shelton, J., Lorenzini, K., Valdez (Alex), G., Williams, T., 2015. Applying dynamic modeling methods to IH 35 through austin: exploring options for addressing future congestion. Int. J. Transp. Sci. Technol. 4 (3), 257–276.

Song, C.J., Zhang, Y.P., Li, Q., 2015. Evolution modeling and simulation on traffic structure in urban areas based on the nested-logit model. In: 15th COTA International Conference of Transportation Professionals, pp. 3948–3959.

Tan, R., Adnan, M., Lee, D., 2015. A new path size formulation in path size logit for route choice modeling in public transport networks. Transp. Res. Rec.: J. Transp. Res. Board 2538, 11–18.

Wan, X., Chen, J., Zheng, M., 2009. The impact of land use and traffic resources supply on commute mode choice. In: 2009 2nd International Conference on Intelligent Computing Technology and Automation. ICICTA 2009, 3, pp. 1010–1013.

Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2011. Statistical and Econometric Methods for Transportation Data Analysis. Chapman and Hall/CRC, Boca Raton, FL.

Xu, C., Wang, W., Li, Z., Yang, C., 2010. Comparative study on drivers' route choice response to travel information at different departure time. In: CAR 2010–2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics 3, pp. 97–100.

Yang, L., Zheng, G., Zhu, X., 2013. Cross-nested logit model for the joint choice of residential location, travel mode, and departure time. Habitat Int. 38, 157–166.