Paper ID: 211

Modelling the Route Choice Behaviour under Stop-&-Go Traffic for different Car Driver Segments

Saxena N.[1,*], Rashidi T.H.[1], Dixit V.V.[1], Waller S.T.[1]

[1] Research Centre for Integrated Transport Innovation, School of Civil and Environmental Engineering, UNSW Australia, Sydney NSW 2052, Australia
[*] Corresponding author: n.saxena@student.unsw.edu.au

Abstract
Stop-&-go (S&G) waves are often a nuisance and lead to increased emissions and safety risks Due to its frequent and annoying nature; drivers tend to avoid travelling on routes with more occurrences of S&G waves. This study evaluates the effect of the number of S&Gs on route choice. An online survey was conducted on a sample of regular car drivers residing in Sydney and its neighbouring suburbs. The collected data was analysed using a latent class choice modelling formulation which provides concise segment specific information making it attractive to policy makers. Results showed that nearly three-quarters of the sample had a negative and significant disutility towards the number of S&Gs. The updated model would facilitate realistic assessment of transportation projects and policy decisions.

Keywords: Stop-&-go traffic, route choice, stated choice experiment, latent class choice model, traffic assignment

1 Introduction

A stop-&-go (S&G) wave, also called traffic oscillation or phantom jam, is a traffic phenomenon that often exists in urban road networks during congestion. Under S&G traffic, vehicles are forced to decelerate and travel at a lower speed, or even come to a halt before resuming their original travel speeds *(1)*. Li et al., (2010) found a cyclic occurrence of S&G waves which alternates between slow (stop) and fast (go) movements *(2)*. This leads to an increase in the fuel consumption, carbon dioxide emissions and safety risks *(3)*. Moreover, numerous studies in the health domain have found that traffic oscillations have a detrimental effect on driver physiology, particularly the cardiovascular measurements like blood pressure and heart rate *(4)*. As drivers have to be more focussed while driving in S&G traffic, it results in an elevated discomfort and frustration level among drivers *(5)*. These factors tend to influence drivers to avoid travelling on routes where S&G conditions are prevalent. Few studies have modelled route choice decisions of car drivers as a function of time spent in S&G traffic *(6) (7)*. However, an effect of the number of S&Gs on route choice hasn't been explored till date, which is believed to differently reflect the level of discomfort than the time spent in these conditions. For example, consider the time spent in S&G traffic on route A as 4 minutes. This duration can be due to an occurrence of 4 stops, each of 1 minute duration or by 8 stops, with 30 seconds duration each. In reality, the latter case is expected to be more cumbersome, causing significant increase in the level of frustration and discomfort among drivers. Thus, the objective of this study was to test the hypothesis: an increase in the number of S&Gs on a route increases its disutility. Additionally, this study also evaluated the willingness to pay (WTP) measures between the number of S&Gs with respect to travel time and cost. This particular trade-off hasn't been estimated so far in the literature, to the best of our knowledge.

This study mainly contributes towards including the occurrences of S&G into existing transportation models which serve as an important tool in planning level decisions and project evaluation. Transportation models comprise a traffic assignment module which routes vehicles such that their generalised cost of travel is minimised. The generalised cost function is usually expressed in terms of travel time and cost. Introducing the number of S&Gs experienced on a route, along with time and cost adds a measure of discomfort to the function. The modified function is expected to provide a more realistic evolution of congestion and travel time patterns. This would further enhance the prediction power of the transportation models in practice and would promote sound project assessment. Moreover, results from the LCCM would provide useful insights to policy makers in understanding the nature of each segment of users and tailor policies according to a target segment.

The paper is organised as follows: Section 2 reviews the previous studies which looked at the effect of S&G traffic on the route choice of drivers and identifies the gaps in research. Section 3 discusses the latent class choice model formulation which would be used for data analysis. Sections 4 and 5 describe the design and data collection respectively, for a D-efficient stated choice experiment to determine the route choice preferences of car drivers. An empirical analysis is conducted in section 6 which gives a better understanding of the available data. Section 7 discusses the results from the model analysis and brings out some key findings. Finally, conclusions, challenges and future research directions are discussed in section 8.

2 Background

Given the cyclic and onerous nature of stop-&-go (S&G) traffic, previous studies have attempted to understand its impact on the route choice behaviour of drivers. One of the first studies in this direction was conducted by Small (1999) *(9)*. The study conducted a stated choice (SC) experiment involving two hypothetical routes, each of which was defined in terms of travel time and the percentage of travel time spent in S&G traffic. Further extensions to this approach were conducted by Hensher (2001a) and Rose et al. (2009), which evaluated the impact of free flow, slowed down and S&G travel time on driver disutility *(6) (7)*. All these studies showed an increased driver disutility as the amount of time spent in S&G traffic became larger. For example, Hensher (2001a) found that car commuters, on average, find each minute of their free flow time equivalent to 2.5 minutes under S&G traffic *(6)*. However, the duration spent in S&G traffic doesn't intuitively give an idea about the number of S&Gs experienced. Thus, it would be interesting to additionally evaluate the impact of the number of S&Gs on route choice.

We now review the data analysis techniques that were used in these studies. Small (1999) used logistic regression on the collected data to evaluate the trade-off between the attributes *(9)*. Although the logit model provides a simple closed form framework, it makes some restrictive assumptions like the IIA property and a uniform or systematic taste preference across individuals *(10)*. This inadequacy led to a widespread application of mixed logit models. A mixed logit model not only estimates the mean parameter effect, but also its standard deviation which explains the preference heterogeneity observed across individuals. An alternate specification of the mixed logit model, the error component logit model, accounts for the panel nature of a dataset by capturing the correlation of the unobserved effects across multiple choice tasks. Hensher (2001a) and Rose et al. (2009) used mixed logit models to quantify the impact of driving in S&G traffic on the route choice of car commuters *(6) (7)*. However, a mixed logit model requires an analyst to provide a priori specification of the mixing distribution for every random coefficient, where each parametric distribution has got its own merits and demerits. Hess et al. (2005) discussed the consequences of selecting an inappropriate mixing distribution on the estimated results *(11)*. Moreover, simply knowing that a parameter is randomly distributed across individuals is of lesser interest to policy makers *(12)*. A latent class choice model addresses the shortcomings of a mixed logit model.

To recap, there exists a research gap when it comes to evaluating the effect of the number of stop-&-gos (S&Gs) experienced on a driver's route choice behaviour. Past studies have looked at expressing a relationship between route choice and the amount of time spent in S&G traffic. However, research from medical sciences show that it's the occurrences of S&G waves that leads to a heightened discomfort than the duration for which it is experienced. Moreover, the statistical technique used for data analysis assumed different mixing distributions to explain the taste heterogeneity, an inappropriate selection of which might lead to an incorrect estimation of parameters. Deriving group specific WTP measures using the LCCM model avoids making parametric distributional assumptions that are associated with the mixed logit framework. Thus, an LCCM provides information on the subgroups of individuals and their choice preferences, which assist the decision makers in coming up with segment specific policies. We now discuss the latent class choice model and its formulation that will be used in this study.

3 Latent class choice model framework

A Latent Class Choice Model (LCCM) is a statistical tool that can reveal the underlying subgroups of individuals from observed multivariate data. The tool, that was first developed in the field of marketing sciences, is a parsimonious technique of clustering the observed choice patterns of individuals into mutually exclusive latent segments *(13)*. Unlike parametric discrete choice models like the mixed logit, LCCM doesn't require any mixing distribution to be assumed upfront. It in turn identifies the latent segments in the population from the observed data like the socio-demographic information. LCCMs have found numerous applications in transportation planning to study the heterogeneity in the mode choice behaviour *(12)* and modality styles of individuals *(8)*.

An LCCM comprises of two components, namely, a class membership model and a discrete choice model. The class membership model expresses the unobserved latent class segments in terms of the available data like the socio-demographic information of individuals. The model can be specified as a multinomial logit (MNL) which estimates a set of coefficients that are the same for individuals within a segment. The choice model, on the other hand, evaluates the probability of observing the response pattern of an individual, conditioned that the individual belongs to a specific latent segment. The response pattern can be a set of choices made by an individual in an SC experiment. The choice model can be specified using different formulations depending upon the nature of available response data. The integrated framework is then run multiple times by progressively increasing the number of latent classes at each run. The optimum number of latent segments is determined based on the three criteria: 1) overall goodness of fit, 2) model parsimony, and 3) behavioural interpretation of latent segments *(8)*. We now discuss the model formulation that will be used in this study.

3.1 Model specification

Consider that a collected dataset for N individuals contain two parts: a cross sectional data on the socio-demographic information and a panel data of choice patterns for every individual. We first discuss the class membership model specification. Assuming the sample comprises C latent class segments, the utility (U_{nc}) for a person n belonging to a latent class c is given by equation 1.

$$U_{nc} = \alpha_c' W_n + \varepsilon_{nc} \tag{1}$$

In this equation, α_c is a vector of parameters that is exclusive to class c. W_n denotes a vector of observed socio-demographic characteristics of n. ε_{nc} represents the idiosyncratic error term and is assumed to follow Gumbel distribution with a variance of $\pi^2/6$. This forms the MNL specification for the class membership model which is given in equation 2.

$$\gamma_{nc} = \frac{\exp(\alpha_c' W_n)}{\sum_{k=1}^{C} \exp(\alpha_k' W_n)} \tag{2}$$

In equation 2, γ_{nc} is the latent class prevalence for individual n being in class c. In order to maintain model identification, one of the latent segments is set as the base category. It means that only $C - 1$ segments can be estimated from a class membership model, with α_c vector for the base category being normalised to zero.

For the choice model, an error component logit specification was used to capture the correlation across multiple choice tasks for individual n. Assume that an individual is presented T choice tasks, each of which comprises J alternatives. The utility ($U_{njt|c}$) that an individual n, belonging to class c, derives from an alternative j in a choice task t is given by equation 3 where X_{njt} is a vector of attributes presented in that choice task for an alternative. β_c is a vector of generic parameters and σ_c is the estimated variance of the error component for every latent segment c. The error component $\xi_{nj|c}$, which is considered to capture the impact of multiple responses by one individual, is assumed to be normally distributed with a mean and variance of 0 and 1 respectively. ε_{njt} is again the idiosyncratic term (like ε_{nc} in equation 1) that follows the Gumbel distribution. Equation 4 gives the logit kernel for evaluating the probability of choosing the alternative in a single choice task.

$$U_{njt|c} = \beta_c' X_{njt} + \sigma_c \xi_{nj|c} + \varepsilon_{njt} \tag{3}$$

$$P_{njt|c} = \int \frac{\exp\left(\beta_c' X_{njt} + \sigma_c \xi_{nj|c}\right)}{\sum_{l=1}^{J} \exp\left(\beta_c' X_{nlt} + \sigma_c \xi_{nl|c}\right)} f\left(\xi_{nj|c}\right) d\xi_{nj|c} \tag{4}$$

Let Y_n be the observed response pattern across T choice tasks for individual n. Then the probability of observing Y_n conditional on latent class c is given by equation 5.

$$P(Y_n|c) = \prod_{t=1}^{T} P_{njt|c} \tag{5}$$

Equation 6 gives the total probability of observing Y_n across C latent segments which is calculated as the expected value of latent class prevalence and its corresponding conditional choice probability.

$$P(Y_n) = \sum_{c=1}^{C} \gamma_{nc} \prod_{t=1}^{T} P_{njt|c} \tag{6}$$

Equation 6 is repeated over all individuals N to give the likelihood function. Equation 7 gives the likelihood function for the LCCM model.

$$L(\alpha, \beta, \sigma) = \prod_{n=1}^{N} \sum_{c=1}^{C} \gamma_{nc} \prod_{t=1}^{T} P_{njt|c} \tag{7}$$

Equation 7 can be maximised to recover the parameter estimates using the numerical optimisation scheme proposed by Broyden-Fletcher-Goldfarb-Shanno (BFGS). Unlike the Newton-Raphson method, the BFGS procedure doesn't require an analytical evaluation of the Hessian matrix, which makes it computationally efficient when the Hessian matrix is unavailable or is too expensive to compute at each iteration.

Upon estimation of equation 7, class specific vector of choice model coefficients β_c was used to evaluate a variety of willingness to pay (WTP) measures. The four attribute effects that were considered in this study were: the total travel time (β_{TT}), time spent in stop-&-go (S&G) (β_{TTS}), number of S&Gs (β_{SnGo}) and vehicle running cost (β_{VRC}). Equations 8-12 give the different WTP measures among the four attributes. Equations (8) and (9) represent the value of travel time savings (VTTS) under overall and S&G traffic respectively. Equation (10) denotes the WTP in dollars associated with an occurrence of a stop. Equations (11) and (12) represent the trade-off between occurrence of S&G against time spent under overall and S&G conditions. The experiment was designed such that the travel time was inclusive of the time spent in S&G. Therefore, equations (9) and (12) have the coefficient for the time spent in S&G as $\beta_{TT} + \beta_{TTS}$. Equation (11) will be a novel contribution from this study towards modifying the existing transportation models in practice to include the occurrences of S&G.

Running cost – Travel time ($/hr)	$= (\beta_{TT}/\beta_{VRC}) \times 60$	(8)
Running cost – Time in stop-&-go ($/hr)	$= \{(\beta_{TT} + \beta_{TTS})/\beta_{VRC}\} \times 60$	(9)
Running cost – No. of stop-&-go ($/stop)	$= \beta_{SnGo}/\beta_{VRC}$	(10)
Travel time – No. of stop-&-go (min/stop)	$= \beta_{SnGo}/\beta_{TT}$	(11)
Time in stop-&-go – No. of stop-&-go (min/stop)	$= \beta_{SnGo}/(\beta_{TT} + \beta_{TTS})$	(12)

4 Stated choice experiment design

Stated choice (SC) experiments are increasingly being used in transportation research for forecasting the impacts of a proposed alternative or a hypothetical policy decision. These methods present a more structured way of selecting choice tasks as compared to a random selection in the fractional factorial method of design *(14)*. These methods provide rich information towards understanding the trade-off among different attributes, which are made by car drivers during route choice. For this study, each choice task in the experiment involved three alternatives: a status quo alternative and two other unlabelled hypothetical routes that were derived from the status quo alternative. Each alternative was defined in terms of four attributes, namely, the travel time, the time spent in S&G traffic, the number of S&Gs experienced and the vehicle's running cost. The vehicle running cost, which included the fuel cost and the maintenance cost was calculated at the rate of 15 cents (AU$0.15) per kilometre *(15)*. The four attributes, along with their corresponding levels (proportions) for designing the two hypothetical unlabelled routes are listed below. The levels for the time spent in S&G and the vehicle running cost were taken from the study by Hensher (2001a) *(6)*.

Attribute name	**Levels**
Travel time (minutes)	- 20%, - 10%, 0, 10%, 20%
Time spent in stop-&-go traffic (minutes)	- 50%, - 25%, 0, 25%, 50%
Number of stop-&-gos experienced	- 50%, - 25%, 0, 25%, 50%
Running cost of vehicle ($)	- 25%, - 12.5%, 0, 12.5%, 25%

This study used a D-efficient design technique to come up with the stated choice tasks. A D-efficient design overcome the limitations of a D-optimal design proposed by *(17)*. A D-efficient design tries to minimise the asymptotic variance covariance (AVC) matrix in order to evaluate significant parameter estimates. Some advantages of using the D-efficient design are: 1) it requires a comparatively smaller sample size than the D-optimal for determining significant parameter estimates *(18)*, 2) it utilises a prior information about the parameter estimates for design which makes it better than the D-optimal design that assumes no prior information *(19)*, and 3) it provides analyst with an additional flexibility to construct an SC experiment in accordance with an econometric model that would later be applied during data analysis. Furthermore, a pivot design experiment was proposed for this study. A pivot design or reference alternative design uses a respondent's knowledge base to derive attribute levels of hypothetical alternatives in an experiment *(21)*. These designs provide a more realistic comparison of alternatives at an individual level. Each segment comprises its own reference alternative which represents the base case for all individuals belonging to that segment. A homogeneous pivot design was generated which presents the same set of choice tasks to every segment *(21)*.

The experiment in this study was designed to target private car commuters residing in Sydney who drive to work in the morning. Key statistics like the number of car drivers and drive time to destination (work) were obtained from the household travel survey (HTS) report for the annual wave 2014-15 *(22)*. Table 1 presents the reference alternatives, for the SC design, that were formed using the available data. The table classifies Sydney's car driver population into six segments on the basis of travel time.

Table 1 show that nearly 60 percent of the population experience a drive time of more than 10 minutes, covering at-least 5 kilometres. Thus, a majority of population is exposed to longer driving periods. The reference alternative for each of the six segments was prepared using the conditions listed below:

i. Travel time attribute was set as the mean of a segment's travel time range
ii. Time spent in stop-&-go (S&G) was kept between 20 to 25 percent of the travel time attribute (as found by *(16)*)
iii. Number of S&Gs experienced was set between 30 to 40 percent of the upper limit of a travel time range in a segment
iv. Vehicle running cost was calculated as a product of cost per kilometre ($0.15) and the travelled distance in each segment

Table 1 Segment wise weightage and reference alternative constructed for the design

Segment	**Travel time range (minutes)**	**Distance (Km)**	**Weight (%)**	**Reference alternative taken for the design**			
				Travel Time (minutes)	**Time in stop-&-go (minutes)**	**Number of stop-&-go**	**Running cost ($)**
1	0-10	2.239	42.4	5	2	4	0.35
2	11-20	6.436	27.2	15	5	7	1
3	21-30	17.219	14.4	25	8	10	2.6
4	31-40	20.260	5.2	35	10	14	3.05
5	41-60	26.147	7.5	50	15	20	3.95
6	> 60	30	3.3	75	20	30	4.5

The choice tasks were generated using the stated choice experimental design package Ngene *(23)*. The reference alternatives discussed in table 1 and prior parameter values of the mean and standard deviation of the attribute effects (taken from *(24)*) were used as inputs for the design. An error component random parameter logit model specification was chosen for the design purpose. The vehicle running cost attribute was treated as non-random for the design. Further, a Bayesian normal distribution was assumed for the mean and standard deviation of all parameters. This improves the accuracy of the design as prior values are not treated as fixed, thus accounting for unobserved variation. Two blocks of 10 choice tasks each were generated for the study. The D-error for both the designs was observed to be 0.7455 which is acceptable as it is less than 1.

5 Data collection

Individuals residing in Sydney who drive to work by car were selected as the target population for this study. The exclusion criteria to the survey included people who: 1) do not reside in Sydney or its neighbouring regions, 2) do not possess a driver's license, and 3) do not drive to work for at-least thrice a week. A respondent satisfying any of these criteria was dropped from the analysis. The data thus collected represents a sample of car drivers who regularly drive by car and have a better perception of stop-&-go (S&G) conditions than the people who drive occasionally.

An online survey instrument was developed for data collection. The survey was administered by a data collection agency, Qualtrics *(25)*, which distributed the survey weblink among its pool of affiliated members. The survey asked a respondent to recollect his/her recent trip to work by car on a weekday morning. The study considered work trips to test the proposed hypothesis.

The current route was constructed using the experienced travel conditions of the respondent and route-1 and 2 are the hypothetical unlabelled alternatives which were pivoted around the current route. A set of rules were coded in the experimental design to avoid situations arising due to unreasonable combination of attributes. The used logics were as follows: 1) setting minimum and maximum values on the revealed travel attributes (10 and 150 minutes for travel time; 2 and 120 minutes for time spent in S&G; 5 and 150 for number of S&Gs and $0.75 and $18 for the vehicle running cost), and 2) bounding the ratio between the travel time and the time spent in S&G between 0.2 and 0.6. In order to avoid confusion during choice tasks, respondents were informed that the total travel time was inclusive of the time spent in S&G traffic. The survey was circulated between Tuesday and Friday every week for around 5 weeks. A total of 288 responses were received from the main survey. The average survey response duration was found to be 9 minutes across all respondents.

6 Empirical analysis
The collected data was cleaned to exclude any incomplete or invalid data. The final effective sample size was 249 respondents which equated to 2490 SC observations. Table 2 gives the socio-demographic statistics of the effective sample.

Table 2 Descriptive statistics on socio-demographic information

Category		**Sample (%)**	**Population (%)**
Gender	Male	52.61	53
	Female	47.39	47
Age	20 years and less	2.01	3.59
	21 to 30 years	20.08	12.98
	31 to 40 years	23.29	16.57
	41 to 50 years	24.10	19.34
	51 to 60 years	21.69	18.23
	60 years and above	8.84	29.28
Income (AU$)	25K and less	13.65	18.48
	25.1K to 50K	24.10	19.94
	50.1K to 75K	25.70	20.82
	75.1K to 125K	26.91	20.82
	125K and above	9.64	19.94
Driving Experience	10 years and less	24.50	Not Available
	10 to 20 years	23.29	
	20 years and above	52.21	
Work Status	Full-Time	60.64	
	Others	39.36	
Occupation	Business	5.62	
	White-collar	62.65	
	Blue-collar	24.50	
	Students	2.81	
	Others	4.42	

The population statistics were obtained from the Sydney HTS report for the year 2012-13 *(22)*. Percentages in the population column correspond to the characteristics of individuals who drive by private vehicle for their commute, which constitute 63.5 percent of the total weekday trips. The survey sample maps well with the population statistics, with up to 5 percent difference between the two proportions for most socio-demographic segments. A deviation of more than 10 percent was observed in the elderly (60 years and above) and high income (125K and above) groups due to their low response rate. Nonetheless, the collected dataset still gives a good representation of the overall population.

Table 3 provides descriptive statistics of the travel related attributes revealed by individuals during the survey. The table reports a mean travel time of 36 minutes and a vehicle running cost of $3.57 for car commuters, with 80 percent of the sample having a drive time and running cost up to 45 minutes and $5.25 respectively. Nearly half of this time is spent driving in stop-&-go (S&G) traffic with around 12 occurrences of stop-&-go experienced on average. In other words, the sampled drivers have to experience congested driving coupled with S&G conditions while travelling to work during morning peaks. Thus, the selected sample is suitable to test the relationship between route choice and the number of S&Gs experienced. The data for the time spent in S&G traffic and the number of S&Gs was found to be over dispersed. These measures are generally hard to perceive compared to travel time and cost, causing few respondents to report an over-estimated value. Nevertheless, the mean of the estimate is still reasonable and can be used for further analysis.

Each respondent was presented with 10 choice tasks in the experiment, which were thoroughly analysed to discern the underlying driver behaviour. Figure 1 shows the choice tasks 3, 5, 6, 8, 9 and 10 belonging to one of the two blocks used in the study.

Table 3 Descriptive statistics for the SC questionnaire

Data	Mean	Std. dev.	Min.	Max.	20th percentile	80th percentile
Travel time (minutes)	36.06	22.89	10	140	20	45
Time in stop-&-go (minutes)	16.47	18.75	0	130	5	25
No. of stop-&-go	12.06	17.25	0	100	3	15
Running cost ($)	3.57	2.86	0.15	17.25	1.5	5.25

Unlike the other choice tasks, these selected tasks demonstrated a clear preference towards an alternative. Within each choice task plot, the four attributes are represented by histograms and the height denotes their levels across the three alternatives. The rectangular boxes at the feet of the histograms give the percentage of respondents selecting a route in a given choice task. The three main types of respondent behaviour that can be identified through visual inspection of the plots are:

i. Individuals who are more inclined towards reducing their travel time and the number of S&G while making route selection. These people do not have a high disutility towards the running cost and are willing to pay a little extra to minimise the travel time and stop-&go. Choice tasks 3, 8 and 10 represent this behaviour.
ii. Drivers who are more likely to select a route having lower running cost and occurrences of S&G. This kind of behaviour was only visible in choice task 5.
iii. Individuals who primarily try to minimise their travel time and running cost. They generally don't consider the occurrence of S&Gs while making their route choice. Choice tasks 6 and 9 depict this kind of behaviour.

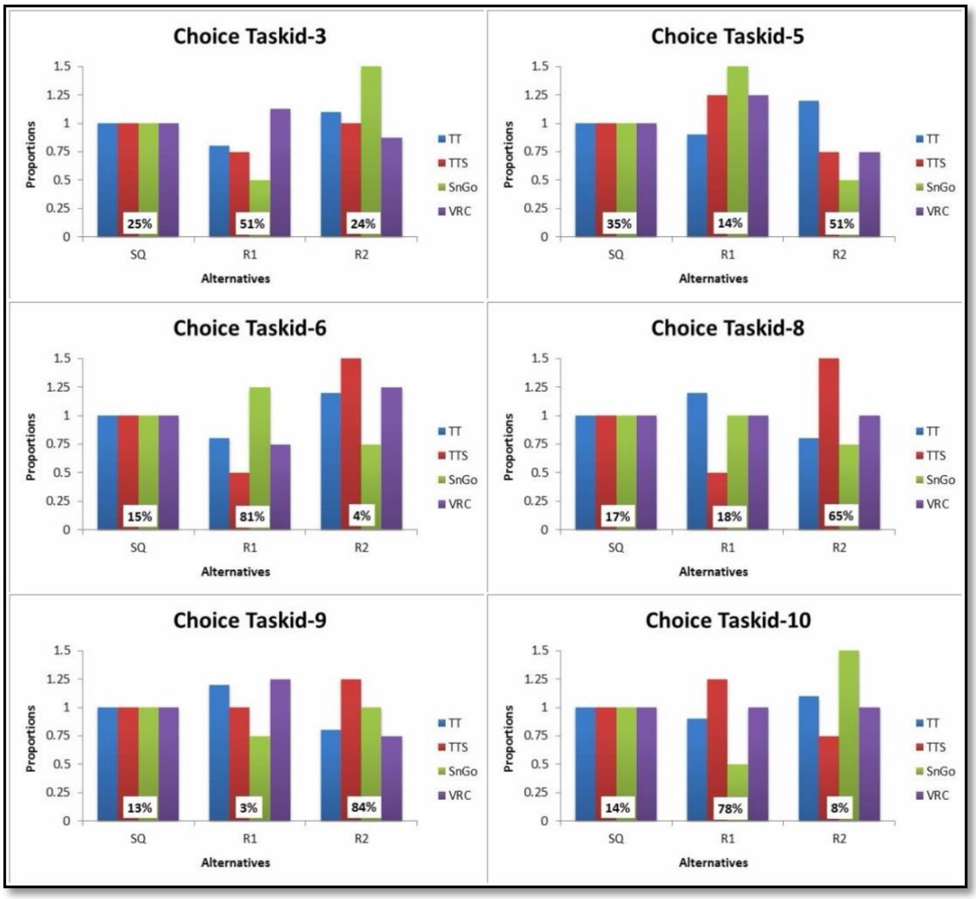

Figure 1 Selected SP choice tasks from block-0 that were presented to respondents. (Percentage of route selection is given in the rectangular boxes)

Identification of these patterns gives us a justification to classify the available dataset on the basis of behavioural interpretation. It would provide useful insights while determining the optimum number and labelling of segments in the latent class analysis.

7 Model results

The latent class choice model (LCCM) specification was coded in Matlab. The choice model consisted of the four attributes from the SC experiment and a normally distributed error component term capturing the correlations across multiple choice tasks. The error components were simulated using 1000 Halton draws from a standard normal distribution to compute the likelihood function (equation 7) *(10)*. The membership model was defined in terms of the socio-demographic characteristics. Of the different socio-economic variables that were tried in the membership model, the model with age, gender and income depicted a superior goodness of fit and model parsimony. The dichotomised age and income variables signify young people (between 20-40 years) versus older and low income (below $25,000) versus high income respectively *(26) (27)*. Once the membership model was identified, the LCCM was re-executed multiple times by incrementing the number of latent segments after each run. From these runs, we finally selected the LCCM with three latent segments for analysis. The selected model satisfied parsimony, least AIC and BIC measures of fit and behavioural interpretation for each of the three latent segments.

Table 4 gives the estimated parameters for the LCCM model, along with the WTP measures for each of the three latent segments. A negative coefficient on age signifies that young people are less likely to be in latent segments 1 and 2. Females are more likely to be in segment 2 than the remaining segments. A negative sign on the income coefficients signifies that individuals in the two subgroups are less likely to be from the low income group. However, the coefficients for income aren't statistically significant at 90 percent confidence which could be due to multicollinearity among categorical explanatory variables. Nonetheless, we still included the income variable in the utility equation because of two reasons: 1) a membership model primarily serves as a prediction model, thus it is acceptable to include covariates which are slightly correlated with one another, and 2) income is a key attribute which is generally used to stratify the population into segments. Interestingly, the model suggests segregation of older and high income people into two different segments.

Results from the choice model show a negative and highly significant parameter values for most of the attributes across the three segments. However, the coefficient for the time spent in stop-&-go (S&G) and the number of stop-&-gos is statistically insignificant and nearly zero for individuals in latent class 2. Moreover, this segment also has a highly negative coefficient for the running cost. A significant error component variance across all segments confirms the presence of the unobserved correlation across multiple choice tasks in the SC experiment. We now discuss some interesting observations from table 4. The identified latent segments can be labelled as follows:

Table 4 Parameter and WTP estimates for a three segment LCCM

Parameters	**Class-1**	**Class-2**	**Class-3**
Class Membership Model			
Constant	0.2888	0.2903	0
Females	0.5744	0.4691 **	0
Age (below 40 years)	-0.8136 ***	-0.6606 ***	0
Income (below 25K)	-0.325	-0.353	0
Choice Model			
Travel time	-0.3948 ***	-0.1913 ***	-0.0413 ***
Time spent in stop-&-go	-0.0951 ***	0.0055	-0.0488 ***
Number of stop-&-gos	-0.0374 ***	-0.0035	-0.1656 ***
Running cost	-0.7085 ***	-3.9599 ***	-0.5231 ***
Sigma (σ)	0.801 ***	0.604 ***	0.4526 ***
Representation in sample (%)	47	22	31
WTP Measures			
Running cost – Travel time ($/hr)	33.43	2.90	4.74
Running cost – Time in stop-&-go ($/hr)	41.49	2.82	10.33
Running cost – No. of stop-&-go ($/stop)	0.05	0.00	0.32
Travel time – No. of stop-&-go (min/stop)	0.09	0.02	4.01
Time in stop-&-go – No. of stop-&-go (min/stop)	0.08	0.02	1.84

*** significant at 99% ** significant at 90%

Class-1 – Patient and highly productive: These constitute 47 percent of the sample population and assign a similar weight on the disutility towards travel time and running cost. In other words, drivers in this segment have a high value of time ($33.43 per hour) which signifies their high productivity. They are in general less perturbed by the discomfort due to the number of S&Gs, which can be seen through a small coefficient value of -0.0374. The trade-off value observed between travel time and the number of S&Gs is also less (0.09 minutes per stop) which signifies 6 seconds of extra travel to reduce the number of S&Gs by one on the travelled route. This segment mainly includes individuals from the older age group who are also in the high income bracket. These people are generally at the pinnacle of their carrier which brings in increased disposable income and additional professional responsibilities on them.

Class-2 – Strict cost minimisers: Comprising 22 percent of the sample population, drivers in this segment have a very high disutility towards the running cost (-3.9599) when compared to travel time (-0.1913) and an insignificant and near zero effect towards S&G traffic. In other words, the drivers in this group pivot their route choice decisions mainly around the cost and are indifferent towards occurrences of S&G. They are generally comfortable on departing early from home to compensate for a route with a longer travel time. A fairly low value of time ($2.90 per hour) explains their purely cost minimising behaviour. Similarly, statistically insignificant trade-off value between travel time and number of S&Gs (0.02 minutes per stop) shows their indifference towards the attribute. This segment also comprises people from the old age and high income groups. Interestingly, females are more likely to be in this segment. This finding is also backed by a study by Srinivasan (2005) which found women to spend lesser on travel than males *(28)*.

Class-3 – Impatient and less productive: At 31 percent of the sample population, drivers in this segment have a high disutility towards the running cost (-0.5231) when compared to travel time (-0.0413), along with a negative and significant disutility towards S&G traffic. They are quite sensitive towards the number of S&Gs experienced, which can be seen through a high disutility coefficient of -0.1656. They generally have a less value of time ($4.74 per hour), but assign a very high weight on travel time to reduce an occurrence of S&G (4.01 minutes per stop). This segment mainly comprises young people who also fall in the low income group, which explains their cost saving attitude. They are generally quite impulsive in nature and easily get frustrated while experiencing alternating cycles of S&G waves.

The latent class choice model discussed above brings out some interesting and important findings. First, results show that nearly 78 percent of the population do consider the number of S&Gs experienced while making a route choice decision. Of this proportion, around 40 percent assign a high weight on reducing the occurrence of S&G. Secondly, the class membership model provides the socio-demographic characteristics of each labelled segment, understanding which can facilitate segment specific policies to be rolled out by planners. We now discuss the policy implications of the findings from this study.

8 Conclusions

Driving in stop-&-go (S&G) traffic leads to increased safety risks and driver distress. In this study, we tested the hypothesis that drivers experience increased disutility from a route with multiple occurrences of S&Gs. A stated choice experiment was conducted on a sample of car drivers in Sydney who regularly drive to work. A latent class choice model (LCCM) was developed to analyse the collected data because of the following advantages: 1) unlike the mixed logit model, LCCM is a non-parametric (or semi-parametric) approach that classifies the underlying behaviour into subgroups using the observed data, and 2) it provides more useful and concise information to policy makers which facilitates formulation of different schemes targeting a particular segment. Results show that nearly three-quarters of the sample has a disutility towards the number of S&Gs experienced. Especially, the analysis found that impatient and less productive users associate a higher disutility with the occurrence of S&G compared to the patient and highly productive group. The study also found a small group of drivers who were more likely to be females and indifferent towards S&G traffic. The outcomes from this study will have a significant contribution towards modifying the existing transportation models and include the number of S&Gs into the assignment algorithms. The resulting transportation model would realistically represent the evolution patterns of congestion, and would promote sound project assessment.

While this work brings out some new findings to add to the existing knowledge, it made an assumption during data collection which is also the study's limitation. The study assumed the number of stop-&-gos revealed by participants as a true measure. Unlike travel time and cost, the number of S&Gs isn't perceived well by participants due to its subjective (latent) nature. Thus, there could be a measurement bias induced into the current results which may lead to incorrect estimation of WTP measures.

Future research works will mainly focus on two aspects. The first task will be to conduct a driving simulator study to understand the route choice behaviour of drivers. The experiments would allow participants to experience realistic traffic scenarios in a virtually controlled environment, thus minimising the measurement bias with respect to the number of stop-&-gos. The second task will be to explore techniques for quantifying the occurrences of stop-&-go waves. Work on both the tasks is currently underway.

References

1. Shott, M. J. *Traffic Oscillations Due To Topology and Route Choice in Elemental Freeway Networks*. UNIVERSITY OF CALIFORNIA, DAVIS, 2011.
2. Li, X., F. Peng, and Y. Ouyang. Measurement and estimation of traffic oscillation properties. *Transportation Research Part B: Methodological*, Vol. 44, No. 1, 2010, pp. 1–14.
3. Zheng, Z., S. Ahn, and C. M. Monsere. Impact of traffic oscillations on freeway crash occurrences. *Accident Analysis & Prevention*, Vol. 42, No. 2, 2010, pp. 626–636.
4. Yang, Y., B. Reimer, B. Mehler, and J. Dobres. A Field Study Assessing Driving Performance, Visual Attention, Heart Rate and Subjective Ratings in Response to Two Types of Cognitive Workload. New York, 2013.
5. Levinson, D., K. Harder, J. Bloomfield, and K. Winiarczyk. Weighting Waiting: Evaluating Perception of In-Vehicle Travel Time Under Moving and Stopped Conditions. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1898, 2004, pp. 61–68.
6. Hensher, D. A. The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications. *Transportation*, Vol. 28, No. 2, pp. 101–118.
7. Rose, J. M., D. A. Hensher, S. Caussade, J. de D. Ortúzar, and R.-C. Jou. Identifying differences in willingness to pay due to dimensionality in stated choice experiments: a cross country analysis. *Journal of Transport Geography*, Vol. 17, No. 1, 2009, pp. 21–29.
8. Vij, A., A. Carrel, and J. L. Walker. Incorporating the influence of latent modal preferences on travel mode choice behavior. *Transportation Research Part A: Policy and Practice*, Vol. 54, 2013, pp. 164–178.
9. Small, K. A. *Valuation of Travel-time Savings and Predictability in Congested Conditions for Highway User-cost Estimation*. Transportation Research Board, 1999.
10. Train, K. E. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge ; New York, 2009.
11. Hess, S., M. Bierlaire, and J. W. Polak. Estimation of value of travel-time savings using mixed logit models. *Transportation Research Part A: Policy and Practice*, Vol. 39, No. 2–3, 2005, pp. 221–236.
12. Hess, S., M. E. Ben-Akiva, D. Gopinath, and J. L. Walker. Taste Heterogeneity, Correlation, and Elasticities in Latent Class Choice Models. Presented at the Transportation Research Board 88th Annual Meeting, 2009.
13. Kamakura, W. A., and G. J. Russell. A Probabilistic Choice Model for Market Segmentation and Elasticity Structure. *Journal of Marketing Research*, Vol. 26, No. 4, 1989, pp. 379–390.
14. Hensher, D. A., J. M. Rose, J. de D. Ortúzar, and L. I. Rizzi. Estimating the willingness to pay and value of risk reduction for car occupants in the road environment. *Transportation Research Part A: Policy and Practice*, Vol. 43, No. 7, 2009, pp. 692–707.
15. RACQ. RACQ – Vehicle Running Cost 2014. http://www.racq.com.au/cars-and-driving/cars/owning-and-maintaining-a-car/car-running-costs. Accessed Oct. 26, 2014.
16. Hensher, D. A. The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications. *Transportation*, Vol. 28, No. 2, 2001, pp. 101–118.
17. Street, D. J., D. S. Bunch, and B. J. Moore. Optimal Designs for 2k Paired Comparison Experiments. *Communications in Statistics - Theory and Methods*, Vol. 30, No. 10, 2001, pp. 2149–2171.
18. Bliemer, M. C. J., and J. M. Rose. Experimental design influences on stated choice outputs: An empirical study in air travel choice. *Transportation Research Part A: Policy and Practice*, Vol. 45, No. 1, 2011, pp. 63–79.
19. Bliemer, M. C., and J. M. Rose. Efficiency and sample size requirements for stated choice studies. *Institute of Transport Studies and Logistics Working Paper*, No. ITLS- WP-05-08, 2005.
20. Bliemer, M. C. J., and J. M. Rose. Construction of experimental designs for mixed logit models allowing for correlation across choice observations. *Transportation Research Part B: Methodological*, Vol. 44, No. 6, 2010, pp. 720–734.
21. Rose, J. M., M. C. J. Bliemer, D. A. Hensher, and A. T. Collins. Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological*, Vol. 42, No. 4, 2008, pp. 395–406.
22. BTS. Bureau of Transport Statistics. http://www.bts.nsw.gov.au/Statistics/Household-Travel-Survey/default.aspx#top. Accessed Feb. 5, 2016.
23. ChoiceMetrics. Ngene 1.1.1 User Manual & Reference Guide, Australia.
24. Saxena, N., V. Dixit, and T. Waller. Influence of Stop-&-Go Conditions on Route Choice Behaviour. Presented at the IATBR 2015 - WINDSOR, 2015.
25. Qualtrics. Qualtrics. https://www.qualtrics.com/. Accessed Feb. 20, 2016.
26. ABS. 6523.0 - Household Income and Wealth, Australia, 2013-14. http://www.abs.gov.au/ausstats/abs@.nsf/mf/6523.0. Accessed Feb. 9, 2016.
27. Erikson, E. H., and J. M. Erikson. *The Life Cycle Completed (Extended Version)*. W. W. Norton, 1998.

28. Srinivasan, S. *Research on Women's Issues in Transportation, Report of a Conference: Technical papers*. Transportation Research Board, 2005.