Australasian Transport Research Forum 2010 Proceedings
29 September – 1 October 2010, Canberra, Australia
Publication website: http://www.patrec.org/atrf.aspx

Freight stakeholders' sensitivities under road user charging: a latent class approach

Sean M. Puckett[1], Simona Rasciute[2]

[1]Institute of Transport and Logistics Studies, Faculty of Economics and Business,

The University of Sydney, Sydney NSW 2006

[2]Department of Economics, Loughborough University, Loughborough, United Kingdom LE11 3TU

Email for correspondence: sean.puckett@sydney.edu.au

Abstract

An understanding of the potential welfare effects of changes in the function of road infrastructure is central to strategic decisions relating to transport policy. This research applies a discrete mixture approach to identify values of travel time savings and reliability gains for freight transport providers and their customers under a hypothetical road user charging system in the Sydney Metropolitan Area. The policy implications from latent class models are compared with those from mixed logit models, highlighting the relative merit of both approaches in transport policy analysis.

This paper confirms the value of taking an alternative approach in representing preference heterogeneity, through the use of the latent class model. By segmenting the sample probabilistically with respect to contextual effects, the latent class model avoids the use of distributional assumptions present in mixed logit models within a more tractable estimation framework. A further empirical development in this research is the use of attribute processing strategies to condition class membership and preference estimates.

1. Introduction

An effective travel demand management instrument offers a stimulus that is designed to influence traveller behaviour in some meaningful way. Under a given set of preferences, travellers will respond to a change in the state of the world in a manner that is (hopefully) consistent with those preferences. Hence, with a reliable knowledge of those preferences, policy makers can feasibly construct tools that lead to a desired change in traveller behaviour. This relationship gains complexity as the structure of traveller preferences becomes more complex; a population with multiple subsets of decision makers, each of which has a distinct set of preferences, is more difficult to influence effectively than a population with homogenous preferences.

Researchers have utilised tools to represent some measure of preference heterogeneity, ranging from basic interaction terms involving socio-demographic characteristics and other contextual effects, to relatively complex econometric models such as the mixed multinomial logit (MMNL) and latent class (LC) models. The primary distinction between these two model types is the nature of the specification of preference heterogeneity. In the MMNL model, the analyst assumes a distribution over which the ensuing individual-specific or choice-set-specific preference estimates are found. In the LC model, the analyst searches for the most representative number of groups, within which decision makers are assumed to hold identical preferences; the probabilistic assignment of respondents to latent classes is aided by the search for the most effective indicators (i.e., contextual effects) of class membership.

This research focuses on the application of the LC model to a study of the preferences of interdependent freight transport stakeholders in the Sydney Metropolitan Area. Although the LC model has been applied in recent travel behaviour research (see Section 2), we know of no existing studies into the merits of the LC model in representing preference heterogeneity in urban goods movement. Rather, the data analysed within this study, which are unique in scope by capturing preferences relating to cost/level-of-service (LOS) trade-offs under variable road user charging, have to this point only been analysed within an MMNL framework.

This paper applies the LC model to a stated preference (SP) experiment of interdependent road freight stakeholders in the Sydney Metropolitan Area, in an effort to gauge the potential to capture preference heterogeneity in urban freight without the distributional assumptions present in the MMNL model, and hence to evaluate the merit of the LC model as an alternative to the MMNL model in urban freight studies. The SP data centre on shippers' and road freight carriers' cost/LOS trade-offs under a hypothetical distance-based road user charging system. The discrete choice models within the analysis identify heterogeneity in key policy outputs, including the value of travel time savings (VTTS) for free-flow and slowed-down travel conditions, and the value of reliability gains. These policy outputs are central to an analysis of the net benefits of changes in the LOS provided under a road pricing system, and offer additional insight to the impacts of changes in the LOS, in general, whether resulting from public policy or changes in service offerings by carriers.

This research takes an additional step within the LC framework, by utilising respondents' stated APSs as inputs in the class membership function, as proposed by Hess and Rose (2007). Respondents in the study were asked to indicate attributes that they ignored or aggregated for each alternative in each choice set presented. The choice to ignore an attribute implies a lack of behavioural significance to the respondent within the context of the particular choice setting. The behavioural link between the choice to ignore an attribute and the underlying sensitivities of the respondent to the attribute level mixes within the choice set is not necessarily directly tangible, and hence is of interest as part of a latent construct identifying preference heterogeneity.

That is, respondents that ignore a particular attribute may not have a total lack of sensitivity to the attribute, but rather the choice to ignore the attribute could be representative of some general set of preferences (e.g., those who ignored slowed-down time within the experiment are more likely to be part of a subgroup that places high importance on on-time reliability). Similarly, the choice to aggregate attributes that share a common metric (e.g., time measures) does not necessarily represent equivalent preferences for each attribute within the aggregation. Rather, the choice to aggregate a particular set of measures may be representative of a propensity to belong to a subgroup with a given set of preferences (e.g., those who combined fuel cost and distance-based charges into one measure are more likely to be time-sensitive).

We present a discussion of the LC model in the following section, clarifying the components of the model and how class membership is determined. Section 3 offers a description of the SP study and the resulting data analysed herein. In Section 4 we present our empirical analysis, in which we compare willingness-to-pay measures from LC models to MMNL models with and without APS information, culminating in an evaluation of policy implications across these competing model structures. The discussion concludes with comments on the merit of LC models and the value of capturing APS information in transport studies.

2. Modelling approaches to freight travel demand

Powerful strides have been made in representing heterogeneous preferences within econometric models. A predominant research tool within the area of discrete choice is the use of the mixed multinomial logit (MMNL) model (also referred to as the random parameter logit model, or simply the mixed logit model), which estimates distributions of preference parameters across a given sample (see Train (2003) for a detailed description of the model). These estimated distributions can be calibrated with respect to the unobserved components of estimated utility functions (i.e., error terms) alone, or by representing preference heterogeneity as a function of contextual effects (e.g., socio-demographic characteristics).

The MMNL model is an important empirical tool in the search for a useful representation of preference heterogeneity, but it is limited by two related shortcomings. Firstly, the distributions of preferences that one obtains from the MMNL model are essentially analytical artifices that are specified by the researcher, in an effort to obtain the best (or most favourable, depending upon one's objectives) model fit. There may be no *a priori* restriction as to the set of reasonable distributional assumptions the researcher could apply to the model, but the choice of distribution can have a marked influence on the behavioural implications that the model identifies. Not only may there be no reason to suspect that the portion of unobserved effects within a model is best reflected through, say, a normal distribution versus a lognormal, uniform or triangular distribution, but there may also be no clear reason why any given constraint on the chosen distributional form (i.e., the relationship between the mean and standard deviation or spread) is a better representation of the underlying preference heterogeneity than another restriction (or none at all).

Secondly, there is no closed-form solution for the MMNL model, due to the intractability of the multi-dimensional integrals present within the modelling structure. This requires the use of simulation methods involving a series of random draws to obtain a solution. The issue here is that the choice of simulation technique and its underlying parameters can influence the resulting model estimates. For example, the analyst could choose to use simple random draws or intelligent draws (e.g., a standard or shuffled Halton sequence) to estimate the distribution of preference estimates. Along with the choice of type of draw, the analyst must specify the number of draws to take, weighing the computational burden of a relatively large number of draws against the relative reliability in the model outputs that would be gained under a larger number of draws. Even the choice of value with which to seed the draw generation process impacts the model outputs under a reasonable number of draws (i.e., seeding the algorithm with "12345" will likely yield different results to seeding the algorithm with "67890").

Ultimately, there are many influences that the researcher imposes directly within the MMNL estimation process that are not linked to the behaviour driving the data being modelled, yet these influences can have fundamental impacts on the resulting behavioural implications of the final model. All else equal, it would be preferable to have the ability to identify meaningful information about preference heterogeneity without imposing such influences in the estimation process. That is, it would be ideal to limit the researcher's impact on the modelling process to the specification of behaviourally meaningful effects to include within the model when representing preference heterogeneity.

One approach that has gained some traction is the use of the LC model, which represents preference heterogeneity by identifying distinct classes of decision makers. Within each class, decision makers' preferences are represented uniformly. This is an important distinction; whilst preference heterogeneity is represented in the MMNL model through the simulation of a continuous mixing function relating to unobserved preferences (i.e., estimating a continuous distribution within the unobserved effects), the LC model represents preference heterogeneity through a discrete mixture (i.e., estimating a discrete set of subgroups with homogeneous preferences). The analyst observes the assignment of each respondent to a particular class only up to a probability (hence the latent nature of the classes), with class membership determined by an analyst-specified model that is calibrated with respect to contextual effects.

The LC model offers a powerful alternative to the MMNL model for two key reasons. Firstly, the identification of latent classes through contextual effects enables the analyst to represent preference heterogeneity in an intuitive, behaviourally meaningful way. That is, converse to the search for a best-fitting distribution of a component of unobserved effects (e.g., testing an unconstrained normal distribution versus a triangular distribution constrained to have a spread equal to twice the mean), the LC model centres on an explicit interaction between preferences and contextual effects (e.g., simultaneously estimating the degree to which preferences for transport level-of-service (LOS) vary across subgroups and the likelihood that respondents with particular characteristics belong to a given subgroup). Secondly, the LC model has a closed-form solution, and hence is not subject to the precision concerns associated with the analyst-specified parameters relating to the random draws within the MMNL model.

These advantages make the LC model an important empirical alternative in the search for useful representations of preference heterogeneity. This is of particular relevance in urban freight, where the heterogeneous preferences of interdependent decision makers (e.g., shippers and freight transport providers) can lead to a wide range of sensitivities to travel demand management policies across individual freight movements and broader freight distribution strategies. That is, efforts to influence travel behaviour, in general, or freight travel behaviour specifically, could lead to divergent changes in freight travel activity due to the heterogeneous preferences underlying freight travel behaviour. Hence, to satisfy policy objectives relating to urban freight travel behaviour, it is critical to capture the nature of preference heterogeneity in urban freight with as much precision as the available data allow. As such, urban freight studies could benefit from utilising an expanded empirical toolkit involving stronger behaviour-based representations of preference heterogeneity.

Some recent transport studies have turned to the LC model for behavioural analysis. Beckman and Goulias (2008) applied a complex LC model to identify preference heterogeneity relating to commuting behaviour. Calibrating their LC model with respect to spatial and socio-economic characteristics were enabled Beckman and Goulias to identify population segments with distinct sensitivities in mode choice, route choice and departure time choice systematically. Wen and Lai (2010) offer an appealing application of the LC model as an alternative to the MMNL model in the identification of behavioural segments in the air travel market. A LC model including socio-demographic characteristics with intuitive links to preference heterogeneity such as age, income, and trip purpose identified two distinct classes of air travellers with different preferences for individual airlines and some level-of-service attributes. Shen (2009) offers direct evidence for the power of LC models in transport studies relative to the MMNL model, by comparing the implications of model results from LC and MMNL models of two stated preference surveys of public transport patronage. Shen found that the LC model offered a significantly better model fit than the MMNL model for both datasets. Whilst the policy outputs (i.e., willingness-to-pay and choice elasticities) from the two models gave similar mean implications, the LC models identified segments with preferences that are highly divergent from those identified in the estimated distributions in the MMNL models.

Most closely related to the research presented in this paper, Hensher and Greene (forthcoming) utilise an inferred attribute processing strategy (APS) approach in a LC model of car drivers in Sydney. In their model, latent classes are specified in terms of *a priori* APS behaviour, with distinct utility expressions in each class constrained to match the assumed APS profile. Despite the lack of observed APS data, this model resulted in an improved fit relative to the MMNL model. Supplementary questions about APS behaviour were included at the end of the survey to gauge the relationship between inferred and respondent-stated

APS behaviour; there was not a strong mapping between inferred and stated APS behaviour, but a major limitation of the technique may have been the absence of any.choice-set-specific APS questions. The analysis presented in Section 4 is calibrated on stated APS information captured for each alternative in each choice set, offering an important distinction from both the inferred and stated behaviour in Hensher and Greene (forthcoming).

The LC model is a semi-parametric extension of the MMNL model, which does not require the researcher to make specific assumptions about the distribution of random parameters across respondents. Rather, preference heterogeneity across respondents is modelled with a discrete distribution. The LC model approximates the unknown distribution of random coefficients by a finite number of mass points; therefore, simulation is not needed in the estimation process (Meijer and Rouwendal, 2006). Respondents are implicitly divided into a number of classes Q, although it is not known which class contains a particular firm. Within this application, shippers' and carriers' behaviour is governed by observable attributes and on latent heterogeneity that varies with factors that are unobserved (Greene and Hensher, 2003).

In the LC model, preference estimates are class specific such that choice observations assigned to a particular class *q* share the same estimated preference parameters β_q corresponding to a vector of independent variables *z* presented in each alternative in the choice set.

The probability that respondent *i* in class *q* chooses alternative *j* is given as:

$$\Pr(Y_{ij} = c | class = q) = \frac{\exp\left(\beta_q z'_{ij}\right)}{\sum_{c=1}^{C} \exp\left(\beta_q z'_{ij}\right)} \qquad (1)$$

Within each class, the choice probabilities are generated by the multinomial logit (MNL) model.

Class membership, however, is not observed; rather, class probabilities are also specified by the MNL form. The probability of respondent *i* belonging to class *q* can be expressed as:

$$H_{iq} = \frac{\exp\left(\theta_q h'_i\right)}{\sum_{q=1}^{Q} \exp\left(\theta_q h'_i\right)}, q = 1, \dots, Q, \theta_Q = 0, \qquad (2)$$

where h_i denotes a set of observed respondent characteristics (e.g., years spent working in their current role, whether costs were aggregated in the choice set). The LC model estimates the probabilities of a respondent belonging to each class, and respondent is assigned to one of the classes on the basis of the largest probability. Due to identification issues, the Qth parameter vector is normalised to zero; as a result, all other coefficients are interpreted relatively to the normalised class.

Combining the conditional choice equation (1) and membership classification equation (2), the joint probability that respondent *i* belongs to class *q* and chooses alternative *j* can be written:

$$P_{ij} = \sum_{q=1}^{Q} H_{iq} P_{ij|q} = \sum_{q=1}^{Q} \left[\frac{\exp\left(\beta_q z'j\right)}{\sum_{c=1}^{C} \exp\left(\beta_q z'_{ij}\right)} \right] \left[\frac{\exp\left(\theta_q h'_i\right)}{\sum_{q=1}^{Q} \exp\left(\theta_q h'_i\right)} \right] \qquad (3)$$

The parameter vectors β_q and θ_q are simultaneously estimated by the maximum likelihood method, and the log likelihood (LL) for the sample is defined as:

$$LL = \sum_{i=1}^{I} \ln P_i = \sum_{i=1}^{I} \ln \left[\sum_{q=1}^{Q} H_{iq} P_{ij|q} \right] \qquad (4)$$

The log likelihood is maximised with respect to the Q structural parameter vectors, β_q, and the Q-1 latent class parameter vectors, θ_q. The issue in the estimation process is the choice of the number of classes, Q, as the comparison of the log likelihoods of models with a different number of classes is not appropriate. While increasing the number of classes increases the fit of the model, it may lead to some coefficients having very large standard errors.

The trade-off between the goodness of fit and the precision of the parameter estimates can be found with the help of information criteria summarised by Shen and Saijo (2007), which could help determine the optimal number of classes, Q. Candidate criteria include the Akaike Information Criterion, Akaike's ρ^2, the Bozdogan Akaike Information Criterion and the Bayesian Information Criterion. A comparison of these or similar metrics was not a dominant concern in this application, however; models with two latent classes were clearly dominant to models with larger numbers of classes, leading to a strong preference for a two-class modelling structure.

3. Data description

A 2004 study of road freight stakeholders in Sydney, Australia centred on capturing information about independent and interdependent preferences of carriers and shippers in the presence of a (hypothetical) distance-based road pricing system. Consistent with other freight studies, the predominant empirical constraints in the study were: (a) a small population from which to draw; (b) a limited research budget; and (c) difficulties in gaining the cooperation of freight stakeholders. A limited number of agents to sample (i.e., freight firms and their clients under contracts involving urban goods movement) requires optimisation on two counts: (1) recruiting a sufficient proportion of the population for the sample and (2) obtaining a sufficient number of choice observations for each respondent. A minimum information group inference (MIGI) experiment was chosen to allow for a relatively larger sample than a stated choice experiment involving direct interaction between sampled group members due to the relative ease of recruiting participants; that is, no temporal coordination of respondents was required (see Hensher and Puckett, 2008 for details of MIGI experiments).

The empirical procedure began by administering the experiment to representatives of freight firms. Centred on a computer-aided personal interview (CAPI) survey with a d-optimal experimental design (discussed in Puckett *et al*., 2007), the MIGI experiment involved three distinct procedures: (1) non-stated-choice questions intended to capture the relevant deliberation attributes and other contextual effects; (2) choice menus corresponding to an interactive (i.e., freight-contract-based) setting; and (3) questions on the attribute processing strategies enacted by respondents within each choice set. After a sampled respondent from a freight firm completed the survey, a client of a freight firm matching the classification offered by the respondent was recruited and given a survey involving the identical series of choice sets faced by the corresponding freight firm.

The data offer a powerful means of gaining behavioural insight under the aforementioned empirical constraints. The use of a CAPI survey allowed the analysis to be centred on real-market data, including revealed preference trip information, and respondent-, firm- and interfirm relationship-specific information. The revealed preference trip information served as an anchor for the attribute levels in corresponding stated preference alternatives. Furthermore, the socio-demographic information not only allowed respondents to refer to a rich representation of the context of the choice task, but also allowed the analysis to evaluate the degree to which preferences change along with contextual effects.

Fundamentally, the choice data gathered from freight transport providers and their customers within Sydney enabled the direct analysis of sensitivities of Sydney freight transport stakeholders to changes in cost/LOS trade-offs under road user charging, which could not be conducted with extant freight travel behaviour data (e.g., regional-level data on vehicle movements by class, in the absence of road user charging). The use of a *d*-efficient experimental design supplemented the power of the choice data by reducing the sample size required to make meaningful inference with respect to trade-offs across the attributes within the choice sets in the questionnaire.

The levels and ranges of the attributes were chosen to reflect a range of coping strategies under a hypothetical distance-based road user charging regime. The reference alternative within each choice set for respondents from freight firms is created using the details specified by the respondent for the recent freight trip. In all cases except for the variable charges, the attribute levels for each of the SC alternatives are pivoted from the levels of the reference alternative, as detailed below. The levels are expressed as deviations from the reference level, which is the exact value specified in the corresponding non-SC questions, unless noted:

Free-flow time: -50%, -25%, 0, +25%, +50%

Slowed-down time: -50%, -25%, 0, +25%, +50%

Waiting time at destination: -50%, -25%, 0, +25%, +50%

Probability of on-time arrival: -50%, -25%, 0, +25%, +50%, with the resulting value rounded to the nearest 5% (*e.g.*, a reference value of 75% reduced by 50% would yield a raw figure of 37.5%, which would be rounded to 40%).

Fuel cost: -50%, -25%, 0, +25%, +50% (representing changes in fuel taxes of -100%, -50%, 0, +50%, +100%)

Distance-based charges: -50%, -25%, 0, +25%, +50% around a base of 50 percent of the fuel cost (i.e., 100 percent of fuel taxes).

Respondents were asked to assume that, for each of the choice sets given, the same goods need to be carried for the same client, subject to the same constraints faced when the reference trip was undertaken. The specific choice task on the initial screen is, 'If your organization and the client had to reach agreement on which alternative to choose, what would be your order of preference among alternatives?' Respondents are asked to provide a choice for every alternative. The available options for each alternative are: (Name of the alternative) is: My 1st choice; My 2nd choice; My 3rd choice; Not acceptable. At least one of the alternatives must be indicated as a first choice, which was not found to be restrictive, given that the reference alternative represents the *status quo*, which was clearly acceptable in the market.

The resulting estimation sample, after controlling for outliers and problematic respondent data, includes 136 transporters and 129 shippers. The transporters response rate was 45% whilst the response rate of shippers was 72%.

4. Empirical results

We now turn to latent class analysis of the data, in which we evaluate the preferences of transporters and shippers under a hypothetical variable, distance-based road user charging system in the Sydney Metropolitan Area. The analysis centres on the joint estimation of utility functions based upon the attribute level mixes presented within the empirical survey and the membership of respondents within a set of classes of unique preferences. That is, the LC models presented in this section focus on the identification of unique classes of preferences with respect to free-flow-, slowed-down- and waiting time in transit, the probability of on-time arrival, fuel cost, distance-based charges and the freight rate charged to the shipper.

Two candidate sets of LC models were estimated for transporters and shippers: standard models that assume that all respondents pay full attention to all information presented to them, and models conditioned on the attribute processing strategies (APSs) indicated by respondents. In the empirical survey analysed here, respondents indicated whether they: (1) ignored a given attribute within an alternative in a choice set; and (2) added up attributes along a common dimension within an alternative a choice set. In the APS-conditioned models presented here, respondents' APS behaviour was represented at the choice-set level, through contextual effects representing the number of times a particular APS was invoked in a given choice set. This enables the APS data to be used as a contextual effect (i.e., external to the content within the alternatives) to condition the preference estimates across the alternatives considered in a choice set. One effect of this level of detail is that respondents can be assigned to multiple latent classes across choice observations, due to APS variation throughout the survey (i.e., one may ignore a given attribute within one choice set, yet pay attention to it in another).

4.1. Transporter models

Table 1 presents the model results for transporters, comparing the results for a basic multinomial logit (MNL) model, our preferred LC model without APS information, and our preferred APS-conditioned LC model:

Table 1: Transporter model results

	MNL		Latent Class				Latent Class-APS			
			Class 1		***Class 2***		***Class 1***		***Class 2***	
Avg. Class Probability	--		.726		.274		.716		.284	
Utility Function Attributes:										
	Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio
Reference Alternative	1.043*	7.63	2.120*	5.90	-0.646	-0.93	2.299*	6.72	-0.615	-0.99
Free-Flow Time	-0.005*	-3.14	0.003	0.78	-0.127^	-2.21	0.001	0.39	-0.095*	-2.58
Slowed-Down Time	-0.013*	-3.29	-0.013#	-1.83	-0.099#	-1.83	-0.014^	-2.09	-0.084*	-2.59
Waiting Time	0.001	0.37	0.018*	2.80	-0.043^	-2.14	0.019*	2.91	-0.036*	-2.71
Prob. of On-Time Arrival	0.023*	3.81	0.055*	3.88	-0.032	-1.25	0.065*	4.74	-0.039	-1.42
Distance-Based Charges	-0.002	-1.33	0.005#	1.72	-0.094#	-1.86	0.005^	1.85	-0.071*	-2.89
Fuel Cost	-0.007*	-4.29	-0.013*	-3.88	-0.004	-0.41	-0.012*	-3.98	-0.004	-0.60
Freight Rate	0.003#	1.82	0.000	0.08	0.050	1.56	0.001	0.28	0.034*	2.40
Class Membership Attributes:										
			Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio
Constant			0.520	1.19			1.988*	3.35		
Years in a Similar Role			0.057*	3.13			0.064*	3.61		
Scheduling by Transporter			0.813*	2.58						
Scheduling by Receiver of Goods			-1.069*	-2.46			-1.215*	-2.80		
Litres of Fuel Used on RP Trip			0.001#	1.72			0.001*	2.67		
Number of Years of Partnership			-0.031*	-2.60			-0.041*	-3.26		
Sender of the Goods Paid for Shipment			-0.599#	1.80						
Routing by Sender of Goods			--	--			-1.022^	-2.10		
Combined All Cost Measures							-0.796*	-3.06		

Table 1 (Continued): Transporter model results

	MNL	Latent Class	Latent Class-APS
Model Fit:			
Log-Likelihood	-499.99	-457.78	-454.65
Akaike Information Criterion	1.868	1.768	1.756
Bayesian Information Criterion	1.931	1.949	1.938
Number of Observations	544	544	544

*-Denotes coefficient is statistically significantly different to zero at the 99 percent confidence level.
^-Denotes coefficient is statistically significantly different to zero at the 95 percent confidence level.
#-Denotes coefficient is statistically significantly different to zero at the 90 percent confidence level.

The MNL model reveals a range of sensible implications relating to the behaviour of transporters. What is missing from the model, however, is information relating to any preference heterogeneity in the sample. We now offer the results from LC models with and without APS information, to highlight the nature of the preference heterogeneity identified through this approach.

Both versions of the LC model presented in Table 1 offer improved data fits relative to the MNL model when evaluated with respect to the log-likelihood ratio or Akaike Information Criterion (AIC). Neither LC model offers a sufficient improvement relative to the increased complexity of the model when using the Bayesian Information Criterion, which carries a relatively strong penalty for increases in model complexity. However, given the clear improvement in terms of both the log-likelihood ratio and AIC, we are confident that both LC models offer improved behavioural implications through modelling preference heterogeneity.

For both LC models, the best model fit was found through a two-class representation; indeed, for many specifications tested, the model would not converge with a larger number of classes. Across the two models in Table 1, the general behavioural implications are the same. A large proportion of the sample (around 73 percent to 27 percent in the non-APS model versus 72 percent to 28 percent in the APS model) appears to hold preferences consistent with the mean effects given in Class 1, in which transporters: hold a very strong preference for the reference alternative; are particularly sensitive to slowed-down travel time and on-time reliability; receive positive utility from waiting time; when faced with alternatives involving distance-based charges, recognise a benefit from the charge (in contrast to the strong resistance to alternatives that include the charge); are relatively sensitive to fuel cost; and are not sensitive to changes in the freight rate.

Conversely, a minority of the sample appears to hold preferences consistent with the mean effects given in Class 2, in which transporters: hold a significant preference for alternatives involving distance-based charges; do not appear sensitive to slowed-down time relative to free-flow time; find waiting time to be a source of disutility; are not highly responsive to changes in on-time reliability; when faced with alternatives involving distance-based charges, identify disutility with the specific level of the charge (in contrast to general support for alternatives that include the charge); are not strongly sensitive to fuel cost; and are sensitive to changes in the freight rate.

The discrepancy in behaviour across these classes is substantial. Around three-fourths of the sample is highly sensitive to the level-of-service of the network, responding to changes in slowed-down travel time, on-time reliability and fuel expenditure for a trip. The general lack of sensitivity Class 1 to changes in the freight rate, along with the positive mean influence of distance-based charges, may indicate a tendency for members of Class 1 to pass along these charges to their customers (i.e., changes in distance-based charges translate to changes in the freight rate). Lastly, the positive utility associated with waiting time within Class 1 implies that many transporters are suitably compensated for time spent waiting at the destination (either directly through financial means or through a positive value of that time being used for purposes such as rest). Overall, whilst members of Class 1 prefer their revealed preference trip, on average, their estimated sensitivities identify a potential opportunity to provide sufficient changes in levels-of-service through distance-based charges to influence the spatial and temporal profile of freight distribution activity.

The behaviour identified in Class 2 suggests that approximately one-fourth of transporters are sensitive to travel time and cost components at an aggregate level. That is, cost/level-of-service trade-offs still matter within Class 2, but it appears to be overall trade-offs between the sum of travel time components and the sum of fuel costs and distance-based charges (offset by changes in the freight rate) that are of paramount concern. These trade-offs appear to have dwarfed the benefits of increased reliability in some alternatives, which is a distinct outcome to the sample-level behaviour identified in the base MNL model (and in Class 1). Members of Class 2 revealed a class-level preference for alternatives involving distance-based charges, and hence this behaviour isolates an important proportion of the sample that may be particularly responsive to a distance-based charging system. The strong sensitivity to the specific level of the charges implies that, whilst around one-fourth of transporters' behaviour in the sample involved a general preference for alternatives involving distance-based charges, the level of the charge could influence these transporters' behaviour more strongly than other transporters.

Differences in the sensitivity of respondents to specific components of cost and time across classes are supported by the inclusion of APS information in the model. In alternative specifications of the APS-conditioned model, improved model fits over the non-APS model were found through the inclusion of an indicator whether the respondent added up time measures or cost measures. The behaviour involved in the choice to aggregate times and costs was similar enough across the two APS choices (i.e., whether to aggregate times and whether to aggregate costs) that the model was not improved by including both APS measures. Rather, the indicator of whether costs were aggregated was included in the APS-conditioned LC model.

The power of including the APS variable in the class membership function was sufficient to offer improved overall model fit whilst removing the explanatory power of two variables in the class membership function of the non-APS model (whether the transporter had influence over the scheduling of the trip, and whether the sender of the goods paid for the shipment). This effect was consistent with the behaviour identified in the model: those who aggregated costs whilst making their choice were more likely to be allocated to Class 2. Once this effect was accounted for, the remaining implications of the class membership function were found to be similar across the two models. That is, respondents were more likely to be allocated to Class 1: the more time they had spent working in a similar professional role; if the receiver of the goods had no influence over the scheduling of the trip; the more fuel was used in the revealed preference trip; the shorter the length of the business relationship between the transporter and the shipper; and if the sender of the goods did not have influence over the route choice for the trip.

The major distinguishing characteristics between the non-APS and APS-conditioned models are the relative sensitivities of members of Class 2 to free-flow and slowed-down time (with the sensitivities approaching parity in the APS model, which is consistent with a tendency to aggregate time measures), and the relative sensitivities of members of Class 2 to monetary attributes; the implications for members of Class 1 are consistent across the two models. Hence, accounting for respondents' stated tendencies to aggregate cost measures appears to strengthen the model's capability both to assign respondents to the appropriate class and to identify the sensitivities of respondents within those classes.

With these differences in mind, we turn to a comparison of willingness-to-pay measures across the models, including values of travel time savings (VTTS) under free-flow and slowed-down conditions, and the value of reliability gains (VRG), with the corresponding values from MMNL models from Puckett and Hensher (2008):

Table 2: Transporter willingness-to-pay

	MNL	Latent Class	Latent Class-APS	Mixed Logit	Mixed Logit-APS
VTTS ($/hr, Free-Flow)	$50.84	*Class 2:* $81.06	*Class 2:* $80.28	Mean: $42.48 Std. Dev.: $22.95	Mean: $42.20 Range: $-55.63-274.53
VTTS **($/hr, Slowed-Down)**	$132.17	*Class 1:* $60.00 *Class 2:* $63.19	*Class 1:* $70.00 *Class 2:* $70.99	Mean: $83.77 Std. Dev.: $8.88	Mean: $51.00 Range: $-55.63-360.48
VRG ($/Percentage Point of Prob. of On-Time Arrival)	$3.90	*Class 1:* $4.23	*Class 1:* $5.42	Mean: $3.54 Std. Dev.: $0.46	Mean: $4.35 Range: $3.79-16.08

Each of the willingness-to-pay values in Table 2 were found by taking the ratio of estimated marginal utility parameters of interest (i.e., free-flow time, slowed-down time, probability of on-time arrival) to an appropriate cost measure. This yields a measure of dollars per unit of measurement. Hourly VTTS measures were obtained by multiplying the appropriate ratios, which give dollars per minute of travel time savings, by 60. Transporters' WTP values from the MNL model were found by establishing the cost-based denominator of the ratio as the weighted average of the estimated parameters for fuel cost and distance-based charges, based upon the proportion of average values for each attribute in the choice sets faced by transporters. For all transporter WTP measures in the LC models, only one cost measure was a statistically-significant source of disutility for any class with a significant sensitivity to the attribute of interest, and hence no weighted average was necessary.

The MNL model suggests a strong discrepancy in the VTTS for free-flow time versus slowed-down time. This is only partially confirmed by the LC models, which show a significant VTTS for slowed-down time for members of Class 1 (who do not show a corresponding sensitivity to free-flow time). Consistent with the implications of the distinction between the two latent classes, those in Class 2 do not demonstrate any increase in disutility in slowed-down travel time relative to free-flow time. Hence, by accounting for preference heterogeneity through the identification of latent classes, not only do the overall estimated sensitivities of transporters to travel time savings appear to be lower than under the MNL model, but the relative premium transporters may be willing to pay to avoid slowed-down time appears to be lower, as well.

Mean sensitivities to reliability are estimated to be roughly equivalent across the MNL and MMNL models. The LC models, however, identify significant sensitivities to on-time reliability for Class 1 alone (over 70 percent of the sample). In the non-APS LC model, the estimated mean sensitivity within Class 1 is within the range of mean values in the MNL and MMNL models. When APS information is accounted for, the estimated mean sensitivity in Class 1 is over one dollar per percentage point increase in reliability higher than in any of the other models; when multiplying this value by the proportion of the sample in Class 1, the value is within the range of mean values in the MNL and MMNL models, confirming sample-level consistency in mean estimates across the model structures.

The impacts of utilising APS information in the estimation process are also distinct between the LC model and the MMNL model when considering travel time savings. The use of APS information to allocate respondents to classes increases the WTP estimates for all measures, and brings the VTTS values for free-flow and slowed-down time closer to one another for Class 2, in which respondents are more likely to have stated that they added cost measures together. In the MMNL model, which involved directly specifying marginal utility parameter estimates to reflect the stated APS information (e.g., a respondent's marginal utility for free-flow time was set to zero if the respondent ignored free-flow time), the difference between mean VTTS measures is also reduced considerably. Hence, both advanced modelling structures involving APS information indicate a potentially large bias when failing to account for both preference heterogeneity and APS behaviour.

4.2. Shipper models

Table 3 presents the model results for shippers, comparing the results for a basic MNL model, our preferred LC model without APS information, and our preferred APS-conditioned LC model:

Table 3: Shipper model results

	MNL		**Latent Class**				**Latent Class-APS**			
			Class 1		***Class 2***		***Class 1***		***Class 2***	
Avg. Class Probability	--		.583		.417		.543		.457	
	Coefficient	***t*-ratio**	**Coefficient**	***t*-ratio**	**Coefficient**	***t*-ratio**	**Coefficient**	***t*-ratio**	**Coefficient**	***t*-ratio**
Utility Function Attributes:										
Reference Alternative	0.902*	9.25	1.409	5.54	0.067	0.21	1.949*	5.85	-0.235	-0.77
Free-Flow Time	-0.008*	-7.08	-0.019*	3.77	0.003	0.88	-0.016*	-2.87	0.000	0.02
Slowed-Down Time	-0.019*	-6.33	-0.034*	-4.29	-0.016*	-2.53	-0.034*	-3.95	-0.021*	-3.35
Waiting Time	-0.006*	-2.97	-0.009#	-1.91	-0.005	-1.07	-0.010#	-1.75	-0.005	-1.17
Prob. of On-Time Arrival	0.060*	10.86	0.000	-0.03	0.195*	4.56	-0.023	-1.23	0.180*	4.70
Distance-Based Charges	-0.001	-1.14	-0.007#	-1.76	-0.005#	-1.83	-0.001	-1.58	-0.005^	-2.05
Fuel Cost	-0.002^	-2.30	-0.018*	-2.75	0.007*	2.42	-0.023*	-2.68	0.007*	2.65
Freight Rate	-0.005*	-5.46	-0.018#	-1.66	-0.003	-1.29	-0.009	-0.93	-0.005*	-2.42

Table 3 (Continued): Shipper model results

	MNL	Latent Class				Latent Class-APS			
		Class 1		***Class 2***		***Class 1***		***Class 2***	
Class Membership Attributes:									
		Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio
Constant		-0.456	-1.39			-0.117	-0.32		
Subsidiary of Larger Firm		-0.465^	2.14			-0.434^	-2.00		
Years in a Similar Role		0.038*	2.65			0.041*	2.96		
Proportion of Business for Transporter		0.016*	2.93			0.013*	2.40		
Hours Available to Meet Delivery		0.006^	2.14						
Number of Delivery Locations		0.057	1.53						
Combined Time Measures						-0.270*	-2.42		
Combined Cost Measures						0.404*	2.57		
Ignored All Waiting Times						-0.860*	-2.75		
Ignored All On-Time Probabilities						1.757*	3.61		
Ignored All Fuel Costs						-1.343^	-1.97		
Ignored All Distance-Based Charges						1.475^	2.09		
Model Fit:									
Log-Likelihood	-841.07	-790.36				-779.89			
Akaike Information Criterion	1.645	1.574				1.560			
Bayesian Information Criterion	1.684	1.680				1.680			
Number of Observations	1032	1032				1032			

*-Denotes coefficient is statistically significantly different to zero at the 99 percent confidence level.
^-Denotes coefficient is statistically significantly different to zero at the 95 percent confidence level.
#-Denotes coefficient is statistically significantly different to zero at the 90 percent confidence level.

Extending the modelling approach to identify LC structures offers a clear improvement in model fit across all measures presented in Table 3. The greatest improvement comes from the choice to specify a LC structure, but further gains are found through the inclusion of APS information (i.e., the log-likelihood function and Akaike Information Criterion improve when APS information is included in the model). However, consistent with the transporter models, the penalty imposed by the Bayesian Information Criterion on additional model parameters indicates that the non-APS model is as suitable a fit as the APS model. Also consistent with the transporter model, the specification of two latent classes gave the best (and sometimes only) fit during the model search process.

Both LC models allocated a majority of observations to Class 1 (58 percent to 42 percent in the non-APS model and 54 to 46 percent in the APS model). In both models, shippers in Class 1 held a strong preference for the reference alternative. With respect to travel conditions, members of Class 1 show a tendency to place a premium on free-flow travel time, and also maintain a disutility for waiting time. Interestingly, members of Class 1 do not appear to find sufficient value in increases in on-time reliability when trading off against other attributes. Members of Class 1 are strongly sensitive to fuel cost relative to distance-based charges. Whilst the statistical significance of shippers' mean sensitivity to the freight rate is different across the two models, both models assign a greater sensitivity to the freight rate to members of Class 1.

In contrast, shippers in Class 2 do not demonstrate any significant preference for or against the reference alternative, and appear to focus their travel time considerations on slowed-down conditions; however, this sensitivity is lower in magnitude than the corresponding mean sensitivity of members of Class 1. Above the focus placed on slowed-down travel time, members of Class 2 place a very high premium on on-time reliability, relative to both the MNL model and members of Class 1 (who do not demonstrate a significant sensitivity to on-time reliability). Shippers in Class 2 reveal a confusing sensitivity to fuel cost, in which members react positively to increases in fuel cost. A plausible explanation may be that Class 2 includes shippers who respond positively to a higher ratio of fuel costs to the freight rate (i.e., lower profit margins for transporters). Members of Class 2 are also more sensitive to distance-based charges relative to other monetary measures than those in Class 1; the sensitivity to the charges is accompanied by a lower sensitivity to the freight rate overall, relative to Class 1.

The major distinction between the latent classes rests in the general attitude toward alternatives involving distance-based charges, and sensitivities to free-flow travel time, on-time reliability and cost measures. Whereas Class 1 includes relatively price-sensitive respondents with a preference for the reference alternative, Class 2 includes respondents who place a high premium on on-time reliability, traded off against slowed-down time and the level of the distance-based charges. Hence, a distance-based charging system is likely to impact shippers in different ways, should the two-class model be representative of preference heterogeneity in the population. Specifically, shippers with preferences consistent with Class 1 would likely trade off changes in the freight rate under a distance-based charging system against overall changes in travel time. This response would be distinct to those with preferences consistent with Class 2, whose response to specific charge levels would be weighed against decreases in slowed-down travel time and increases in on-time reliability.

In both the transporter and shipper models, the impacts of including APS information are generally isolated to one class. Whilst these effects were isolated to transporters who tended to include those who aggregated measures (i.e., Class 2 of the transporter models), these effects are more complex in the shipper model; estimated preferences of shippers with respect to distance-based charges and the freight rate in Class 1 are distinct once accounting for a range of APS information. The addition of APS information in the shipper model results in a much lower estimated sensitivity to distance-based charges and the freight rate within Class 1. That is, accounting for APS effects results in a strong downward adjustment in the estimated sensitivity of shippers represented in Class 1 with respect to both the direct transport costs they face and the distance-based charges that transporters could face.

Just as in the transporter model, the inclusion of APS information not only improves the explanatory power of the model, but also crowds out the explanatory power offered by two non-APS contextual effects. In this case, hours available to meet delivery requirements and the number of delivery locations (both of which increase the probability that an observation is assigned to Class 1) are no longer statistically significant after APS information is added to the model. The remaining non-APS contextual effects maintain their general impacts on the class membership model after including APS information; status as a subsidiary of a larger firm decreases the likelihood that an observation is assigned to Class 1, whilst experience working in a similar role and a high proportion of the transporter’s activity being devoted to the business relationship increase the likelihood of being assigned to Class 1.

Consistent with the transporter model, combining cost measures influences class membership (with a greater likelihood of being assigned to Class 1). However, combining time measures has the opposite impact, increasing the likelihood of being assigned to Class 2. Furthermore, four choices whether to ignore individual attributes influence class membership in the shipper model. These indicators are generally consistent with the distinct utility functions across the two classes. The choice to combine time measures increases the likelihood of being assigned to Class 2, which involves a weaker distinction across time measures. Ignoring waiting time also increases the probability of being assigned to Class 2, where waiting time is not a significant influence on utility. Similarly, ignoring on-time probabilities increases the likelihood of being assigned to Class 1, where on-time reliability is not a significant factor. The choice to ignore fuel cost increases the likelihood of being assigned to Class 2, in which shippers appear to associate a positive utility with fuel cost. The choice to ignore distance-based charges increases the likelihood of being assigned to Class 1, in which shippers demonstrate no significant sensitivity to the charges. The odd indicator out is the choice to combine cost measures, which increases the probability of being assigned to Class 1, in which fuel cost is the dominant cost-related influence on utility.

Table 4 highlights the range of estimates of shippers’ values of reliability gains (VRG) across the models, with the values from the MMNL models in Puckett and Hensher (2008) offered for comparison:

Table 4: Shipper value of reliability gains

Model	**VRG ($/Percentage Point of Prob. of On-Time Arrival)**
MNL	$12.00
LC (Class 2)	$10.83
LC-APS (Class 2)	$36.00
MMNL (Mean)	$10.32
MMNL (Standard Deviation)	$1.94
MMNL-APS (Mean)	$11.76
MMNL-APS (Standard Deviation)	$3.72

Shippers’ VRG estimates were found by dividing the estimated marginal utilities of on-time reliability by the estimated marginal disutilities of the freight rate (i.e., the cost paid by shippers), multiplied by -1 (to represent a beneficial change). The MNL model yields an estimated mean VRG of $12 per percentage point increase in the probability of on-time arrival, which is roughly consistent with the mean values in both MMNL models. The LC models identify a segment of shippers (between approximately 42 and 46 percent of the sample) with significant sensitivities to reliability improvements. Consistent with the estimated VRG values for transporters, the non-APS LC model yields a mean sensitivity within the range of mean estimates in the MNL and MMNL models. Likewise, as with transporters, by accounting for APS information yields an estimated mean sensitivity within the segment that is higher than in the other models, and which, when weighting the estimate by the proportion of the sample within the segment, approaches the mean values in the other models.

For both the LC and MMNL models, accounting for APS information results in an increase in the estimated sensitivities of shippers to reliability improvements. However, the magnitude of the increase is considerably larger in the LC model.

5. Concluding remarks

Our empirical results have confirmed the findings in recent transport studies (e.g., Shen, 2009; Hensher and Greene, forthcoming) that the LC model can offer superior explanatory power relative to the MMNL model in the identification of significant preference heterogeneity. Whilst considerably more study is required to draw strong conclusions about the relative merits of the LC model, the results highlight the potential to observe meaningful variation in preferences of transport stakeholders with a reduction of analyst-imposed *a priori* constraints on the nature of preference heterogeneity in a given sample. In this application, we have identified important variation in the preferences of buyers and sellers of road freight transport services in the Sydney Metropolitan Area, as shown by the existence of two distinct classes of behaviour for transporters and shippers, with a corresponding distinction in policy outputs (i.e., willingness-to-pay) in each group. For both transporters and shippers, one subset of the sample demonstrated a clear sensitivity to gains in the reliability of freight transport services as a function of the levels-of-service and costs across alternatives. Likewise, the two classes for both transporters and shippers demonstrated divergent preferences with respect to distance-based charges and the costs associated with each alternative.

In general, the estimated sample-level sensitivities were consistent between the LC models shown here and the MMNL results from Puckett and Hensher (2008). Hence, population-level estimates of welfare benefits resulting from LC and MMNL models may not be significantly different in a given application. The major difference between the two approaches in this vein would be the specific distributional implications. Whilst the MMNL results for the data analysed herein suggest the presence of some population-level mean and mode willingness-to-pay for savings in free-flow and slowed-down travel time and increases in reliability, the LC results focus on the contrast in segment-level average sensitivities for these attributes.

These two approaches give two quite separate forms of insight into potential population-level sensitivities; the MMNL model can give a clean distribution of preferences, whilst the LC model gives point (mean) estimates for discrete segments of the population. What is most important to debate is whether the analyst-induced analytical distributional assumptions that generate the estimated distributions in the MMNL model offer any meaningful improvement over the unknown variation around the estimated segment-level mean sensitivities in the LC model. The relative model fits in this study, along with Shen (2009) and Hensher and Greene (forthcoming) suggest that the distributional content of the MMNL model may not offer any analytical improvement behaviourally.

A further encouraging result is the improvement in model fit found when accounting for respondents' stated attribute processing strategies. The inclusion of APS information not only improved model fit, but also led to some differences in willingness-to-pay estimates relative to the standard model. Furthermore, the inclusion of APS information offered sufficient behavioural information to reduce the number of other contextual effects required to optimise the estimation of the class membership models. Importantly, the types of sensitivities implied by the estimates for each class were generally consistent with the stated APS information, supporting the validity of the APS information stated in each choice set. This choice-set-specific APS data collection approach is an important level of detail distinguishing this analysis from the inferred APS approach in Hensher and Greene (forthcoming), which compared the allocation of respondents to pre-designed APS-based classes with the global (i.e., not specific to each choice set) stated APS information given by respondents at the end of the survey in their study. Given that respondents may demonstrate a range of APS behaviour across choice sets with differing trade-offs, capturing APS data at the choice-set level is an appealing alternative.

References

Beckman, J.D. and Goulias, K.G. 2008, "Immigration, residential location, car ownership, and commuting behaviour: a multivariate latent class analysis from California", *Transportation*, Vol. 35, pp. 655-671.

Greene, W.H. and Hensher, D.A. 2003, "A latent class model for discrete choice analysis: contrasts with mixed logit", *Transportation Research Part B*, Vol. 37, pp. 681-698.

Hensher, D.A. and Greene, W.H. (forthcoming) "Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification", *Empirical Economics*, DOI:10.1007/s100181-009-0310-x.

Hensher, D.A. and Puckett, S.M. 2008, "Power, concession and agreement in freight distribution chains subject to distance-based user charges", *International Journal of Logistics: Research and Applications*, Vol. 11, No. 2, pp. 81-100.

Hess, S. and Rose, J.M. 2007, "A latent class approach to modelling heterogeneous information processing strategies in SP studies", invited paper presented at the Oslo Workshop on Valuation Methods in Transport Planning, Oslo.

Meijer, E. and Rouwendal, J. 2006, "Measuring welfare effects in models with random coefficients", *Journal of Applied Econometrics*, Vol. 21, pp. 227-244.

Puckett, S.M., Hensher, D.A., Rose, J.M. and Collins, A. 2007, "Design and development of a stated choice experiment in a two-agent setting: interactions between buyers and sellers of urban freight services", *Transportation*, Vol. 34, No. 4, pp. 429-451.

Puckett, S.M. and Hensher, D.A. 2008, "The role of attribute processing strategies in estimating the preferences of road freight stakeholders under variable road user charges", *Transportation Research Part E*, Vol. 44, No. 3, pp. 379-395.

Shen, J. 2009, "Latent class model or mixed logit model? A comparison by transport mode choice data", *Applied Economics*, Vol. 41, No. 22, pp. 2915-2924.

Shen, J. and Saijo, T. 2007, "Does energy efficiency label alter consumers' purchase decision? A latent class approach on Shanghai data", Osaka School of International Public Policy, Osaka University, Toyonaka, Osaka.

Train, K. 2003, *Discrete Choice Methods with Simulation*. Cambridge University Press: Cambridge, United Kingdom.

Wen, C-H and Lai, S-C 2010, "Latent class models of international air carrier choice", *Transportation Research Part E*, Vol. 46, pp. 211-221.