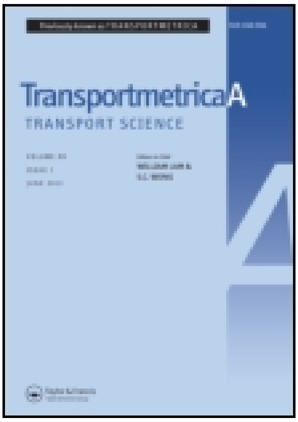

Transportmetrica A: Transport Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ttra21

A latent class generalised nested logit model and its application to modelling carrier choice with market segmentation

Chieh-Hua Wen [a] , Wan-Wen Huang [a] , Chiang Fu [b] & Pei-Yu Chou [a]

[a] Department of Transportation Technology and Management , Feng Chia University , 100, Wenhwa Road, Seatwen , Taichung , 40724 , Taiwan

[b] Institute of Traffic and Transportation, National Chiao Tung University , 4F, 118, Sec. 1, Chung-Hsiao W. Road, Taipei , 10012 , Taiwan

Accepted author version posted online: 04 Jan 2012.Published online: 21 Feb 2012.

Transportmetrica A: Transport Science, 2013
Vol. 9, No. 8, 675–694, http://dx.doi.org/10.1080/18128602.2011.653998

A latent class generalised nested logit model and its application to modelling carrier choice with market segmentation

Chieh-Hua Wen[a]*, Wan-Wen Huang[a], Chiang Fu[b] and Pei-Yu Chou[a]

[a]*Department of Transportation Technology and Management, Feng Chia University, 100, Wenhwa Road, Seatwen, Taichung 40724, Taiwan;* [b]*Institute of Traffic and Transportation, National Chiao Tung University, 4F, 118, Sec. 1, Chung-Hsiao W. Road, Taipei 10012, Taiwan*

(*Received 25 October 2009; final version received 24 December 2011*)

This article develops a new latent class (LC) model with a generalised nested logit (GNL) formulation to enhance the methodology of market segmentation analysis. The standard generalised nested logit (or the cross-nested logit model) is a special case of the latent class generalised nested logit model (LCGNL) that accounts for heterogeneity in individuals' preferences by a number of segments and simultaneously identifies segment sizes and individual profiles. In addition, the LCGNL model allows flexible substitution patterns among alternatives. This extends the standard LC model with the multinomial logit formulation, such that the independence from irrelevant alternatives property does not hold within segments. The proposed model was used to identify potential segments of travellers' preferences towards air and bus carriers. The estimation results of the LCGNL models indicate that degrees of competition vary across carriers and that differential sensitivity in preference parameters exists between segments. The LCGNL model outperforms the other models; therefore, it is a better approach for analysing carrier choice behaviour.

Keywords: latent class; logit; discrete choice; market segmentation; carrier

1. Introduction

The traditional discrete choice model applies the identical values of the unknown parameters to all individual decision makers. Such a model that does not account for taste heterogeneity may fail to represent the underlying choice behaviour of all decision makers. To account for the individual heterogeneity, discrete choice analysis often incorporates a market segmentation scheme to capture variations in taste parameters across individual decision makers. Various segmentation approaches have been used in the literature to analyse travel choices. The standard and most popular segmentation approach specifies the variable(s) to be used to segment the population into a finite number of segments and typically adopts socioeconomic factors (e.g. income or car ownership) or trip patterns (e.g. trip purpose and trip length) to create segments (Pels *et al*. 2001, Loo *et al*. 2006, Bekhor

*Corresponding author. Email: chwen@fcu.edu.tw

and Elgar 2007, Loo 2008, Rastogi and Krishna Rao 2009). Given pre-determined market segments, the discrete choice model is further applied to estimate segment-specific parameters for each segment. An alternative approach does not determine the number of segments in advance, and statistical classification techniques such as cluster analysis (e.g. Psaraki and Abacoumkin 2002, Outwater *et al*. 2004) and automatic interaction detection (e.g. Badoe and Miller 1998) are used to establish segments and profiles of individuals in each segment. Such an approach must perform segmentation and choice modelling in a two-stage and sequential procedure.

Originally developed by Kamakura and Russell (1989) for brand choice, the latent class (LC) choice model allows simultaneously performing segmentation and choice modelling that jointly determine segment-specific parameters, individual profiles of each segment and segment sizes. Both the LC choice model and mixed logit model (Revelt and Train 1998, McFadden and Train 2000) can allow preference heterogeneity across individuals. However, these two models accommodate variations of taste parameters in different ways. The mixed logit model uses continuous probability distributions to account for preference heterogeneity across individuals, while the LC choice model uses discrete representation of systematic variations in taste parameters across segments. The advantage of the LC model is that it can explicitly reveal the number, sizes and characteristics of segments. The LC model has recently become widely noted in transportation research for the studies of mode choice (Bhat 1997, Arunotayanun and Polak 2011), road type choice (Greene and Hensher 2003), location choice (Walker and Li 2007), airline choice (Teichert *et al*. 2008, Wen and Lai 2010), traffic accident types (Depaire *et al*. 2008) and household car ownership choice (Zhang *et al*. 2009).

The LC choice model defines the unconditional choice probability as a mixture of two probability functions: the conditional probability of choosing an alternative, given that an individual belongs to a segment, and the membership probability of the individual in the segment. The standard LC choice model uses the multinomial logit (MNL) formulation (referred to the LCMNL model) for the conditional choice probability. The MNL model has been noted for the restrictive assumptions (i.e. independent and identical error distributions) that lead to the property of independence from irrelevant alternatives (IIA). Consequently, the use of the MNL and LCMNL models is inadequate when the data contradicts the IIA property. To address the weakness of the LC model with the MNL formulation, Kamakura *et al*. (1996), Bodapati and Gupta (2004) and Swait (2003) developed the LC nested logit (LCNL) model. The LCNL model relaxes the IIA property by allowing the error terms of alternatives in a common nest to be correlated. However, the LCNL model still suffers from the disadvantages of the NL model. The NL model imposes equal covariance among alternatives within the same nest. In addition, the error term of an alternative in a nest is uncorrelated with those of alternatives in other nests because each alternative allows being a member of only one nest.

The restriction on the covariance structure of the NL model can be relaxed by allowing each alternative to be a member of more than one nest. Mostly on the basis of the generalised extreme value (GEV) model (McFadden 1978), several types of discrete choice models offer a more flexible structure for the covariance among alternatives than the NL model can provide. These advanced GEV models include the ordered GEV (Small 1987), the paired combinatorial logit (PCL) (Chu 1989, Koppelman and Wen 2000), the cross-NL (CNL) (Vovsha 1997), the MNL-ordered GEV (Bhat 1998), the generalised NL (GNL)

(Wen and Koppelman 2001), network GEV (Bierlaire 2002, Daly and Bierlaire 2006, Newman 2008) and spatial correlated logit (Bhat and Guo 2004). Each of these GEV models defines a slightly different generating function, which leads to a different model structure. All the models have closed-form formulations for choice probabilities and allow differential covariance among alternatives (Bekhor and Prashker 2008).

The LC model accommodates taste heterogeneity just as the mixed logit model does, but it can explicitly identify the sizes and characteristics of segments, and thus is more adequate for the purpose of segmentation. However, the existent LC models (i.e. LCMNL and LCNL) do not allow a flexible structure for the covariance among alternatives. On the other hand, the recently developed GEV-based models (e.g. CNL and GNL) allow for a more general covariance structure, but they can only be used for segmentation and choice modelling within a two-stage procedure in the current literature. Therefore, integrating the advanced GEV models into the LC framework would allow heterogeneity in individuals' preferences and a flexible structure for the covariance among alternatives, while simultaneously identifying the segment sizes and profiles. The objective of this research is to present a new LC model formulated on the GNL model. The latent class generalised nested logit (LCGNL) model allows taste heterogeneity in utility function parameters just as the standard LC model does, but the LCGNL retains the advantage of the GNL model with a flexible covariance structure. The LCGNL model includes the LC models with the MNL and NL formulations as special cases. Although this article adopts the GNL model formulation, development of new LC models with other advanced GEV formulations is fairly straightforward. Our development of the LCGNL model provides a methodological advancement to the existing body of published research.

Two data sources (air and bus carrier choices) were used to empirically illustrate the applicability of the proposed model. This study estimated a variety of LC segmentation models to identify travellers' preferences towards the service attributes of carriers and to specify the individual characteristics of each segment. Carriers can use the findings to establish successful operational and marketing tactics.

2. Model structure

2.1. *LC model*

The LC model uses discrete variations in taste parameters to accommodate preference heterogeneity across individuals. The LC model rests on the assumption that there exist S segments (where S is to be determined) that are relatively homogeneous within each segment but heterogeneous across segments. Given a finite and fixed number of segments, the LC model estimates segment-specific sets of parameters. The likelihood that individuals belong to a segment is a probabilistic function that can depend on individual characteristics (Gupta and Chintagunta 1994). The intrinsic alternative preferences and the sensitivity to observed measures relating to alternative attributes are identical across individuals within each segment, but are heterogeneous across segments.

The LC model assumes that the utility function of an individual n for any alternative j, given that it belongs to segment s, can be expressed as

$$U_{nj|s} = \beta_s' X_{nj} + \varepsilon_{nj|s}, \tag{1}$$

where X_{nj} is a vector of observable attributes, β_s is a vector of segment-specific parameters to be estimated and $\varepsilon_{nj|s}$ captures a random error of the utility function. The probability that an individual n chooses an alternative j is given by

$$P_n(j) = \sum_{s=1}^{S} P_n(j|s) \cdot M_n(s), \tag{2}$$

where $P_n(j|s)$ is the choice probability of an alternative j by individual n within segment s ($s = 1, 2, \ldots, S$). The segment membership function $M_n(s)$ is the probability that individual n belongs to segment s. Both $P_n(j|s)$ and $M_n(s)$ have the MNL formulation for the standard LC model; Equation (2) can be rewritten as

$$P_n(j) = \sum_{s=1}^{S} \left[\frac{\exp(\beta_s' X_{nj})}{\sum_{j' \in C_n} \exp(\beta_s' X_{nj'})} \cdot \frac{\exp(\gamma_s' Z_n)}{\sum_{s=1}^{S} \exp(\gamma_s' Z_n)} \right], \tag{3}$$

where C_n is the choice set for individual n, including alternative j, Z_n is a vector of membership function variables that contain individual characteristics and γ_s is a vector of parameters for segment s. For identification, membership function coefficients of one segment are set to zero. The segment membership function may be a set of segment-specific constants if none of the individual characteristics are incorporated. Under these circumstances, the membership function probabilities are functions of S segment-specific constants; one constant is set to zero and $(S-1)$ constants are estimated.

Estimations of complex segment membership functions require inclusion of individual characteristics (e.g. individual socioeconomic and trip variables). After parameter estimates have been identified, for each pairing of individuals and segments, one can calculate the probability that the individual belongs to that segment. The size of a specific segment is the average of individual membership probabilities. Furthermore, if individual characteristics are included in each segment membership function, individual profiles of each segment can be obtained.

2.2. *GNL model*

The GNL model allows more differential covariance among alternatives than many other GEV models, such as MNL, NL, PCL and CNL. The initial CNL model was developed by Vovsha (1997); it is a special case of the GNL model. The extended CNL model is a modification of the original CNL model formulation and is very similar to the GNL model (e.g. Ben-Akiva and Bielaire 1999, Papola 2004, Bierlaire 2006, Bajwa *et al.* 2008). This article adopts Wen and Koppelman's (2001) GNL model formulation that has the following probability function:

$$P_n(j) = \sum_{m} \left\{ \frac{[\alpha_{jm} \exp(\beta' X_{nj})]^{1/\mu_m}}{\sum_{j' \in N_m} [\alpha_{j'm} \exp(\beta' X_{nj'})]^{1/\mu_m}} \cdot \frac{\left[\sum_{j' \in N_m} [\alpha_{j'm} \exp(\beta' X_{nj'})]^{1/\mu_m}\right]^{\mu_m}}{\sum_{m'} \left[\sum_{j' \in N_{m'}} [\alpha_{j'm'} \exp(\beta' X_{nj'})]^{1/\mu_{m'}}\right]^{\mu_{m'}}} \right\}, \tag{4}$$

where N_m is the set of all alternatives included in nest m. β is a vector of utility function parameters to be estimated. α_{jm} is the allocation parameter which characterises the share of alternative j assigned to nest m; α_{jm} must satisfy the condition, $0 \leq \alpha_{jm} \leq 1$, and the additional condition $\sum_m \alpha_{jm} = 1, \forall j$. μ_m is the logsum or dissimilarity parameter for nest m; μ_m must comply with the condition, $0 < \mu_m \leq 1$, for the GNL model to be consistent with random utility maximisation. In the case of $\mu_m = 1$ for all m, no correlation (independence) exists among any alternatives in any nests.

2.3. *LCGNL model*

Most applications of the LC model still use the MNL formulation, which exhibits the IIA property within each segment. Such LC models result in biased estimates and incorrect predictions if similarities exist among groups of alternatives within segments. This study combines the advantages of the LC and GNL models into the development of a new LCGNL model. The probability formulation of the LCGNL model is an extension of Equation (2), as follows:

$$P_n(j) = \sum_{s=1}^{S} \left[\sum_m P_n(j\,|m, s) \cdot P_n(m|s) \right] \cdot M_n(s). \tag{5}$$

Within the segment s, the conditional probabilities that individual n chooses alternative j in nest m are:

$$P_n(j\,|m, s) = \frac{[\alpha^s_{jm} \exp(\beta'_s X_{nj})]^{1/\mu^s_m}}{\sum_{j' \in N_m} [\alpha^s_{j'm} \exp(\beta'_s X_{nj'})]^{1/\mu^s_m}}. \tag{6}$$

where α^s_{jm} is the allocation parameter which represents the portion of alternative j assigned to nest m for segment s; α^s_{jm} must comply with the condition, $0 \leq \alpha^s_{jm} \leq 1$, and the additional condition $\sum_m \alpha^s_{jm} = 1$, $\forall j, s$. μ^s_m is the logsum parameter for nest m and segment s; μ^s_m must satisfy the condition, $0 < \mu^s_m \leq 1$, for the LCGNL model to be consistent with maximization of random utilities.

Correspondingly, within the segment s, the marginal probabilities that individual n is in nest m can be written as

$$P_n(m|s) = \frac{\left[\sum_{j' \in N_m} [\alpha^s_{j'm} \exp(\beta'_s X_{nj})]^{1/\mu^s_m}\right]^{\mu^s_m}}{\sum_{m'} \left[\sum_{j' \in N_{m'}} [\alpha^s_{j'm'} \exp(\beta'_s X_{nj'})]^{1/\mu^s_{m'}}\right]^{\mu^s_{m'}}}. \tag{7}$$

The segment membership function $M_n(s)$ has the same form as the MNL:

$$M_n(s) = \frac{\exp(\gamma'_s Z_n)}{\sum_{s=1}^{S} \exp(\gamma'_s Z_n)}. \tag{8}$$

The GNL model can be regarded as a special case in which the LCGNL model has only one segment. Moreover, the LCGNL model includes several LC models as special types. They can be determined approximately through the restrictions on allocation and logsum parameters. Table 1 illustrates the differences between LCGNL and other LC models. The LCMNL model without grouping alternatives sets all logsum parameters equal to one, and sets the allocation parameter to one for each alternative. The LCNL model allows alternatives to be grouped into nests with distinct logsum parameters, such that the allocation parameter is equal to one for each alternative because each alternative appears

Table 1. Structural comparison of different LC models.

Type	Choice probabilities within segment
LCMNL	$P_n(j\|s) = \dfrac{\exp(\beta_s' X_{nj})}{\sum_{j' \in C_n} \exp(\beta_s' X_{nj'})}$
LCNL	$P_n(j\|s) = P_n(j\|s,m) \cdot P_n(m\|s)$ $= \dfrac{[\exp(\beta_s' X_{nj})]^{1/\mu_m^s}}{\sum_{j' \in N_m} [\exp(\beta_s' X_{nj'})]^{1/\mu_m^s}} \cdot \dfrac{\left[\sum_{j' \in N_m} [\exp(\beta_s' X_{nj'})]^{1/\mu_m^s}\right]^{\mu_m^s}}{\sum_{m'} \left[\sum_{j' \in N_{m'}} [\exp(\beta_s' X_{nj'})]^{1/\mu_{m'}^s}\right]^{\mu_{m'}^s}}$
LCCNL	$P_n(j\|s) = \sum_m P_n(j\|s,m) \cdot P_n(m\|s)$ $= \sum_m \left\{ \dfrac{[\alpha_{jm}^s \exp(\beta_s' X_{nj})]^{1/\mu^s}}{\sum_{j' \in N_m} [\alpha_{j'm}^s \exp(\beta_s' X_{nj'})]^{1/\mu^s}} \cdot \dfrac{\left[\sum_{j' \in N_m} [\alpha_{j'm}^s \exp(\beta_s' X_{nj'})]^{1/\mu^s}\right]^{\mu^s}}{\sum_{m'} \left[\sum_{j' \in N_{m'}} [\alpha_{j'm'}^s \exp(\beta_s' X_{nj'})]^{1/\mu^s}\right]^{\mu^s}} \right\}$
LCGNL	$P_n(j\|s) = \sum_m P_n(j\|s,m) \cdot P_n(m\|s)$ $= \sum_m \left\{ \dfrac{[\alpha_{jm}^s \exp(\beta_s' X_{nj})]^{1/\mu_m^s}}{\sum_{j' \in N_m} [\alpha_{j'm}^s \exp(\beta_s' X_{nj'})]^{1/\mu_m^s}} \cdot \dfrac{\left[\sum_{j' \in N_m} [\alpha_{j'm}^s \exp(\beta_s' X_{nj'})]^{1/\mu_m^s}\right]^{\mu_m^s}}{\sum_{m'} \left[\sum_{j' \in N_{m'}} [\alpha_{j'm'}^s \exp(\beta_s' X_{nj'})]^{1/\mu_{m'}^s}\right]^{\mu_{m'}^s}} \right\}$

in only one nest. The LC cross-nested logit (LCCNL) model allows each alternative to appear in different nests just as the LCGNL model does, but the LCCNL model imposes equality constraints on the logsum parameters of all nests, if following the initial CNL formulation proposed by Vovsha (1997).

2.4. *Estimation approach*

The estimation procedure of the LCGNL model requires simultaneous calibration of utility function, logsum, allocation and membership parameters. In this study, the constrained maximum likelihood module of GAUSS statistical software (Aptech Systems 2008) was used to impose constraints on parameters; this estimation approach is similar to that of the traditional GNL. As Bhat (1997) suggested, maximisation of the log-likelihood function using the common routines in the LC model can be computationally unstable. Initially, it may be necessary to rely on simple models and computationally tractable values. Estimations of the MNL model can provide the estimates of various utility function specifications and identify the important explanatory variables affecting the choice. Based on the preferred MNL model and its parameter coefficients as the starting values, the NL and GNL models can be estimated. In particular, various estimation procedures to search for a preferred GNL model structure are discussed in Wen and Koppelman (2001).

If any behaviourally interpretable nested structures are identified when estimating standard NL and GNL models, analysts can estimate the LCNL and LCGNL models.

Given a pre-specified number of segments, estimations of LCNL and LCGNL models can be implemented in two ways. Analysts can estimate the LCMNL model using the MNL coefficients as starting values. Subsequently, the LCNL and LCGNL models with additional parameters (e.g. logsum) can be estimated using the LCMNL parameter coefficients. Alternatively, analysts can use the NL and GNL coefficients as starting values to estimate the LCNL and LCGNL models, respectively.

To search a preferred specification for a particular LC model (e.g. LCMNL model), various models with different numbers of segments should be estimated and measures of goodness of fits should be evaluated. Determination of the best number of segments for a specific LC model involves an assessment of measures such as the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and adjusted likelihood ratio. If LL is the value of the log-likelihood function at convergence, k is the number of parameters in the model and N is the sample size, $\text{AIC} = -2LL + 2k$ and $\text{BIC} = -2LL + k\ln(N)$. The adjusted likelihood ratio is similar to likelihood ratio index, but it accounts for the number of parameters estimated in the model. An LC model with lower values of AIC and BIC is preferred over the LC models with higher values. The adjusted likelihood ratio lies between 0 and 1, and the value close to 1 indicates the model fits the data well. In addition, segments must be interpretable (Bucklin and Gupta 1992). Although additional segments may improve the model fit, it may not be worthwhile if further insight is not supplied or segment interpretability is difficult. The LCGNL model includes several LC models as special cases. Given the number of segments, the likelihood ratio test can be used to test the LCGNL model versus other LC models (e.g. a three-segment LCGNL vs. a three-segment LCNL).

3. Results and discussion

3.1. *Empirical data*

The first data used to illustrate the models came from studies by Wen *et al*. (2009) and Wen and Lai (2010). The stated preference survey was conducted face-to-face with air travellers at the international terminal at Taiwan Taoyuan International Airport in 2008. The respondents had taken at least one international flight on the Taipei–Tokyo route. Each respondent was asked to evaluate three hypothetical scenarios in the experiments. The survey data included 322 valid respondents.

The survey included stated preference questionnaires that asked air travellers to choose one of four air carriers: China Airlines (CAL), EVA Airways (EVA), Japan Asia Airways (JAL) and All Nippon Airways (ANA), all offering service from Taipei to Tokyo. In the stated preference questionnaires, the attributes used to describe air carriers included quantitative and qualitative attributes (Table 2). Air fare, preferred departure time (i.e. the time difference between preferred and actual flight departure time), flight frequency and punctuality (i.e. on-time performance) were quantitative variables, while the rest of the variables were qualitative in nature. Each qualitative attribute had three levels (e.g. cabin crew service had three levels, including very unfriendly, friendly enough and very friendly).

The second data used to demonstrate the proposed models came from Wen *et al.* (2011). The intercity bus travel survey data included 717 valid respondents. Each respondent was asked to evaluate nine hypothetical scenarios in the stated preference experiments, with a total of 6453 observations. The survey included stated preference

Table 2. Stated preference experiments of airline and bus choice datasets.

Data	Attributes	Types	Levels
Airline	Fare	Quantitative	High, medium, low
	Preferred departure time	Quantitative	No time difference, early 120 min, late 120 min
	Frequency	Quantitative	1,3,5 flights/day
	Punctuality	Quantitative	On time, 30 min late, 60 min late
	Airport check-in service	Qualitative	Very friendly, friendly enough, very unfriendly
	Cabin crew service	Qualitative	Very friendly, friendly enough, very unfriendly
	In-flight seat space	Qualitative	Very comfortable, comfortable enough, very uncomfortable
Bus	Fare	Quantitative	High, medium, low
	Travel time	Quantitative	120, 150, 180 min
	Frequency	Quantitative	Every 10, 25, 40 min
	Punctuality	Quantitative	On time, 10 min late, 20 min late
	Personnel attitude	Qualitative	Very friendly, friendly enough, very unfriendly
	Driver behaviour	Qualitative	Very safe, safe enough, very unsafe
	On-board comfort	Qualitative	Very comfortable, comfortable enough, very uncomfortable

questionnaires that asked intercity bus travellers to choose one of four carrier alternatives: bus carriers A–D. The attributes used to describe bus carriers included quantitative and qualitative attributes (Table 2). Bus fare, frequency, travel time and punctuality were the quantitative variables, while the rest of the variables were qualitative, of which there were three levels (e.g. the levels for personnel attitude include very unfriendly, friendly enough and very friendly).

3.2. *Airline choice models*

3.2.1. *Results of LCNL models*

Three airline-specific constants were included; ANA was selected as the reference carrier. Each qualitative variable (i.e. check-in service, in-flight seat space and cabin crew service) had three levels. Using the lowest level (e.g. 'very unfriendly') as the base category, each variable was modified to create two 0–1 dummy variables for the medium and highest levels. Each qualitative variable had two parametric coefficients, both of which were expected to have positive signs. A large estimated value was expected for the highest of the three qualitative levels.

Estimation of the MNL model can identify the important explanatory variables that have been reported in Wen and Lai (2010). A variety of standard NL models can be estimated using the same utility specification as the preferred MNL model. Two behaviourally interpretable NL models were identified: EVA–ANA nested and JAL–ANA nested (Wen *et al.* 2009). Two logsum parameters were within the 0–1 range and were statistically different from one at the 0.05 level of significance (t-value > 1.96). JAL and ANA are Japanese national carriers and, as expected, they may share some common attributes. EVA and ANA are strategic partners under a code-sharing agreement. Such cooperation causes air travellers to perceive EVA and ANA as similar. Surprisingly, although CAL and EVA are Taiwanese carriers, the logsum parameter of a nested model

Table 3. The LCNL model results: airline choice.

	LCNL Model 1		LCNL Model 2	
Variable	Segment 1 Coeff. (t-value)	Segment 2 Coeff. (t-value)	Segment 1 Coeff. (t-value)	Segment 2 Coeff. (t-value)
CAL constant	0.087 (0.50)	−5.024 (−8.50)	0.027 (0.18)	−4.383 (−11.79)
EVA constant	0.778 (6.56)	−8.573 (−8.23)	0.624 (4.95)	−7.224 (−6.92)
JAL constant	−0.155 (−0.98)	14.325 (7.57)	0.039 (0.43)	12.672 (6.44)
Airfare	−0.068 (−4.57)	−5.093 (−8.20)	−0.054 (−3.56)	−4.429 (−7.23)
Preferred departure time	−0.052 (−5.78)	−0.272 (−4.40)	−0.050 (−6.35)	−0.258 (−5.60)
Flight frequency	0.004 (0.16)	2.269 (6.22)	0.005 (0.24)	1.969 (5.32)
Punctuality	0.131 (5.91)	−0.095 (−1.05)	0.119 (5.66)	−0.073 (−1.14)
Check-in (friendly enough)	0.392 (3.37)	−0.863 (−1.33)	0.248 (1.97)	−0.718 (−1.40)
Check-in (very friendly)	1.187 (7.42)	−2.685 (−5.00)	1.041 (6.66)	−2.107 (−4.74)
Seat (comfortable enough)	0.293 (2.42)	−2.642 (−4.39)	0.275 (2.47)	−2.173 (−5.26)
Seat (very comfortable)	0.564 (4.31)	9.369 (6.33)	0.577 (5.29)	8.353 (5.90)
Cabin (friendly enough)	0.225 (2.19)	5.810 (5.04)	0.218 (2.45)	4.896 (5.01)
Cabin (very friendly)	0.529 (4.55)	6.035 (5.83)	0.493 (4.44)	5.059 (4.92)
Logsum (t-value vs. 1)				
EVA and ANA nest	0.819 (1.50)	0.819 (1.50)		
JAL and ANA nest			0.573 (3.72)	0.573 (3.72)
Segment size	82%	18%	81%	19%
Final log-likelihood	−1066.349		−1062.260	
Likelihood ratio	0.2037		0.2068	
Adjusted likelihood ratio	0.1828		0.1859	

for these two carriers was greater than one, indicating CAL and EVA do not have common unobserved attributes (e.g. the safety records and company images of CAL and EVA are quite distinct). These standard NL models indicate that competitive structures exist among airlines and that degrees of competition vary across airlines.

Given a pre-specified number of segments, analysts can estimate LCMNL model using the MNL coefficients as starting values. The proper number of segments was evaluated with the values of AIC, BIC, and adjusted likelihood ratio. LCMNL models with two or more segments were attempted, but as the number of segments increased to three, the parameter estimates became very unstable and difficult to interpret due to occurrence of small segment probabilities. This phenomenon also occurred when the LCNL and LCGNL models were estimated. Consequently, the two-segment model was adopted for the rest of the analysis.

Table 3 reports estimation results of two LCNL models using the LCMNL coefficients as starting values. The initial model estimations allowed different logsum parameters for two segments. However, our experience suggests that logsum estimates in a single segment can often be insignificant or out of range. Consequently, the estimation results tended to give one segment an NL structure and the other an MNL structure. On the other hand, when the LCNL model imposed equality constraints on the logsum parameters for both segments with the same model structure, the estimation was computationally feasible, and the logsum parameters were more stable.

LCNL Model 1 (EVA and ANA in a nest) had a logsum parameter of 0.819 that was within the 0–1 range, but statistically insignificant at the 0.05 level of significance. LCNL

Model 2 (JAL and ANA in a nest) had a logsum parameter of 0.573 that was within the 0–1 range, and statistically different from one at the 0.05 level of significance. Two LCNL models outperformed the corresponding LCMNL model in terms of the goodness-of-fit measures and the likelihood ratio test at the 0.05 level of significance. LCNL Model 2, with the adjusted likelihood ratio index 0.1859, had a log-likelihood value that was better than that of LCNL Model 1.

Both LCNL models had a small number of air travellers (19%) allocated to Segment 2, and approximately 81% of air travellers in Segment 1 that was a relatively large market. In Segment 1, the coefficient estimates of ticket price and preferred departure time had negative signs, as expected. This indicates that any increase in the value of any of these variables would decrease the utility of the specific carrier and thus reduce the probability that the airline would be chosen. Similarly, positive estimates were associated with flight frequency and punctuality. For each qualitative variable, two coefficient estimates had the positive and expected signs, and a large value was obtained for the highest of the three qualitative levels. A series of *t*-tests showed that, with the exception of flight frequency, these parameters were significantly different from zero, at the 0.05 level.

Air travellers in Segment 2 were more sensitive to airfare, preferred departure time, flight frequency and cabin crew service. Segment 2 yielded negative and counterintuitive coefficients on punctuality, check-in service (friendly enough), check-in service (very friendly) and in-flight seat space (comfortable enough), which indicates that air travellers in this segment were not as concerned about these services as were travellers in Segment 1. It, however, does not imply that the passengers prefer uncomfortable seats or poor check-in service; as each qualitative attribute had three levels in the stated preference experiments, it is possible that the passengers selected an airline that had uncomfortable seats or poor service. Besides, this may have been partially caused by the small number of air travellers allocated to Segment 2. One may argue that the insignificance or counterintuition of some coefficients may impact the model performance for prediction and market share estimation. Because this research does not produce predictions of airline market shares, such problems are irrelevant.

3.2.2. *Results of LCGNL models*

The standard GNL models were estimated using the same utility specification as the preferred MNL model and the nested structures obtained from the two standard NL models. The preferred GNL model had both EVA–ANA and JAL–ANA nests as depicted in Figure 1; this indicated more flexible airline competition than the standard NL models have indicated. Table 4 presents the estimates of the preferred two-segment LCGNL models with and without socioeconomic and trip variables in the segment memberships. As with the LCNL model, the LCGNL Model 1 imposed equality constraints on logsum and allocation parameters for all segments. The nested structure of the LCGNL Model 1 collapsed to that of the GNL model, with two logsum and ANA allocation estimates for each segment. The JAL–ANA logsum estimate was statistically different from one at the 0.05 level of significance, whereas the EVA–ANA logsum estimate was not statistically significant at the 0.05 level, just as with the LCNL model. However, the significance of the logsum parameters was consistent between the LCGNL and LCNL models. Based on the likelihood ratio tests, LCGNL Model 1 outperformed two LCNL models at the 0.10 level

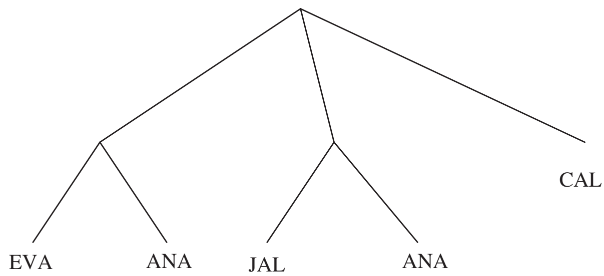

Figure 1. The nested structure of the GNL model (airline choice).

Table 4. The LCGNL model results: airline choice.

Variables	LCGNL Model 1 Segment 1 Coeff. (*t*-value)	LCGNL Model 1 Segment 2 Coeff. (*t*-value)	LCGNL Model 2 Segment 1 Coeff. (*t*-value)	LCGNL Model 2 Segment 2 Coeff. (*t*-value)
CAL constant	−0.149 (−0.95)	−5.173 (−3.23)	−0.139 (−0.88)	−6.735 (−2.36)
EVA constant	0.487 (3.72)	−9.346 (−3.24)	0.499 (3.77)	−13.086 (−2.15)
JAL constant	−0.175 (−1.17)	16.002 (3.50)	−0.170 (−1.12)	21.515 (2.46)
Airfare	−0.051 (−4.07)	−5.567 (−3.62)	−0.053 (−4.14)	−7.331 (−2.63)
Preferred departure time	−0.049 (−6.64)	−0.315 (−3.17)	−0.049 (−6.64)	−0.382 (−2.66)
Flight frequency	0.009 (0.40)	2.573 (3.01)	0.009 (0.38)	3.532 (2.30)
Punctuality	0.104 (5.35)	−0.069 (−0.66)	0.105 (5.39)	−0.062 (−0.40)
Check-in (friendly enough)	0.209 (1.91)	−0.655 (−0.46)	0.217 (1.97)	−0.557 (−0.25)
Check-in (very friendly)	0.953 (7.33)	−2.798 (−2.55)	0.965 (7.38)	−3.919 (−2.13)
Seat (comfortable enough)	0.256 (2.54)	−2.927 (−2.69)	0.262 (2.58)	−4.000 (−1.98)
Seat (very comfortable)	0.528 (5.28)	10.445 (3.49)	0.535 (5.30)	13.989 (2.52)
Cabin (friendly enough)	0.237 (2.70)	6.321 (3.25)	0.239 (2.70)	8.236 (2.67)
Cabin (very friendly)	0.516 (5.29)	6.455 (3.18)	0.513 (5.24)	8.594 (2.43)
Logsum (*t*-value vs. 1)				
EVA and ANA nest	0.780 (1.48)	0.780 (1.48)	0.786 (1.41)	0.786 (1.41)
JAL and ANA nest	0.365 (6.38)	0.365 (6.38)	0.374 (6.12)	0.374 (6.12)
Allocation				
EVA and ANA nest				
EVA	1.000 (–)	1.000 (–)	1.000 (–)	1.000 (–)
ANA	0.472 (3.38)	0.472 (3.38)	0.470 (3.25)	0.470 (3.25)
ANA and JAL nest				
JAL	1.000 (–)	1.000 (–)	1.000 (–)	1.000 (–)
ANA	0.528 (3.77)	0.528 (3.77)	0.530 (3.66)	0.530 (3.66)
CAL	1.000 (–)	1.000 (–)	1.000 (–)	1.000 (–)
Segment membership				
Constant			2.804 (4.69)	
Number of trips			−0.611 (−1.70)	
Income			−1.046 (−2.04)	
Segment size	81%	19%	90%	10%
Final log-likelihood	−1059.294		−1054.577	
Likelihood ratio	0.2090		0.2125	
Adjusted likelihood ratio	0.1866		0.1879	

of significance, but it did not outperform the LCNL Model 2 at the 0.05 level ($\chi^2 = 5.93 < 5.99$) due to the insignificant logsum estimate.

The LCGNL Model 2 had segment membership function that incorporated individual socioeconomic and trip characteristics; its results are shown in Table 4. Segment 2 was used as a basis for comparison, and all membership coefficients in Segment 2 were normalised to zero. The membership coefficient estimates in Segment 1 were interpreted relative to Segment 2. The results indicate that negative coefficients were related to frequent international air travellers (number of trips to overseas) and high-income (personal monthly income > NT$20,000) travellers. Most travellers in Segment 1 were low-income and had less frequent travels; in addition, these passengers were more likely to choose EVA, which was indicated by the positive air carrier constant. Air travellers in Segment 2 were high-income and had more frequent travels; these passengers preferred to use JAL or ANA, and were very unlikely to use EVA. The results elucidate travellers' behavioural profiles. The LCGNL model with socioeconomic and trip variables in the segment memberships had the best goodness-of-fit of all the LC models in terms of the log-likelihood, likelihood ratio index and adjusted likelihood ratio index. According to the likelihood ratio test at the 0.05 level of significance, the LCGNL model with socioeconomic and trip variables in the segment memberships outperformed the corresponding LCGNL model without these variables ($\chi^2 = 9.43 > 5.99$), as well as the LCNL Model 2 ($\chi^2 = 15.37 > 9.49$).

3.3. *Bus choice model*

3.3.1. *Results of MNL and GNL models*

The estimation results of the MNL and GNL models for the bus carrier choice are reported in Table 5. One of the alternatives, bus carrier D, was set as the base (i.e. alternative specific constant was zero). For each of the qualitative variables, two 0–1 dummy variables (1 indicates the designated level, and 0 otherwise) were created for the medium and highest levels, using the lowest level as the base category. Taking driver behaviour as an example, two 0–1 dummy variables for 'safe enough' and 'very safe' were created, while 'very unsafe' was used as the base. Dividing bus fare by personal monthly income was found to enhance the goodness-of-fit, indicating that the fare sensitivity of bus travellers decreases with an increase in personal income. All attributes had expected signs and were significantly different from zero, providing evidence that passengers' choices of intercity bus carriers were associated with these variables.

This study further used the GNL model to account for similarity among intercity bus alternatives. The preferred GNL model consisted of two bus carrier nests, A–B and A–C, as depicted in Figure 2, indicating a high substitution pattern between bus carriers. As bus carriers A and B both offer lower pricing for passengers, they can be grouped into the same nest. In addition, bus carriers A and C were in the same nest because they provide a similar level of service for passengers. The logsum estimates for two nests were within the 0–1 range and were significantly different from one. Using the likelihood ratio test at the 0.05 significance level, the GNL model outperformed the MNL model as well as two corresponding NL models (bus carriers A and B in a single nest and bus carriers A and C in a single nest).

Table 5. The MNL and GNL model results: bus choice.

Variables	MNL model Coeff. (*t*-value)	GNL model Coeff. (*t*-value)
Bus A constant	0.116 (3.87)	0.304 (5.50)
Bus B constant	0.536 (2.84)	0.396 (6.52)
Bus C constant	0.748 (0.04)	0.545 (10.00)
Fare	–0.568 (–11.12)	–0.568 (–12.30)
Travel time	–0.939 (–16.07)	–0.781 (–13.67)
Frequency	0.099 (7.35)	0.087 (7.47)
Punctuality	0.239 (9.06)	0.208 (11.96)
Comfort (comfortable enough)	0.537 (12.97)	0.433 (11.07)
Comfort (very comfortable)	0.927 (22.92)	0.770 (17.40)
Driver (safe enough)	0.390 (9.40)	0.323 (8.71)
Driver (very safe)	0.831 (20.57)	0.698 (16.66)
Attitude (friendly enough)	0.407 (10.25)	0.323 (8.85)
Attitude (very friendly)	0.597 (15.03)	0.483 (12.40)
Logsum (*t*-value vs. 1)		
Bus A and B nest		0.509 (5.47)
Bus A and C nest		0.706 (5.68)
Allocation		
Bus A and B nest		
Bus A		0.377 (5.41)
Bus B		1.000 (–)
Bus A and C nest		
Bus A		0.623 (8.94)
Bus C		1.000 (–)
Bus D		1.000 (–)
Final log-likelihood	–7753.425	–7733.992
Likelihood ratio	0.1333	0.1355
Adjusted likelihood ratio	0.1318	0.1337

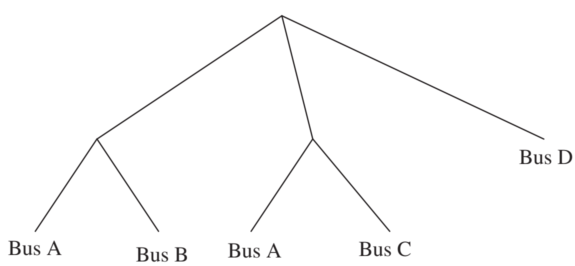

Figure 2. The nested structure of the GNL model (bus choice).

3.3.2. *Results of LC models*

Various LCMNL models with different numbers of segments (i.e. two, three, four and five segments) were estimated. Table 6 presents the fit measures, such as log-likelihood, adjusted likelihood ratio, AIC and BIC, of determining the number of segments. The LCMNL is identical to the MNL when the number of segments is one. The five-segment solution had the lowest BIC and AIC values as well as the largest log-likelihood and likelihood ratio. However, the size of one segment was relatively small (2%), which

Table 6. Goodness-of-fit measures of LCMNL models: bus choice.

Segment	Number of parameters	Log-likelihood	Adjusted likelihood ratio	AIC	BIC	Segment size
1 (MNL)	13	−7753.425	0.1318	15533	15621	1.0
2	27	−7361.428	0.1741	14777	14960	0.20, 0.80
3	41	−7209.694	0.1895	14501	14779	0.22, 0.27, 0.51
4	55	−7011.505	0.2101	14133	14505	0.21, 0.23, 0.27, 0.29
5	69	−6920.772	0.2186	13980	14447	0.02, 0.20, 0.22, 0.27, 0.29

caused the difficulty of interpreting the meaning of that segment. Therefore, the four-segment solution was preferred because it had the lower BIC and AIC values as well as the larger log-likelihood and likelihood ratio than the two- and three-segment results. The four-segment solution was also adopted for the LCNL and LCGNL models.

Since the interpretable nested structures have been identified by the NL and GNL models, the LCNL and LCGNL models were estimated. In order to simplify estimation procedures, the LCGNL model was directly estimated. The initial estimations of the LCGNL models imposed equality constraints on the logsum and allocation parameters across segments; however, a logsum estimate (bus carriers A and C in a single nest) became statistically insignificant. Therefore, the final estimations of LCGNL models allowed for differential logsum parameters across segments and imposed restrictions on the logsum estimates within a reasonable range.

Table 7 provides the estimates of the preferred four-segment LCGNL models without socioeconomic and trip variables in the segment memberships. The nested structure of Segments 2 and 3 collapsed to that of the GNL model, with two logsum and bus A allocation estimates for these two segments. The nested structure of Segments 1 and 4 collapsed to that of the NL model with bus carriers A and B in a single nest. The logsum and allocation estimates were statistically significant at the 0.05 level of significance. Based on the likelihood ratio tests at the 0.05 level, the LCGNL model outperformed two corresponding LCNL models: the LCNL model with bus carriers A and B in a single nest ($\chi^2 = 16.23 > 5.99$) and the LCNL model with bus carriers A and C in a single nest ($\chi^2 = 26.98 > 5.99$).

Segment 1 was the largest one with 33% of the total. The coefficient of bus fare was statistically insignificant in this segment. Bus travellers in this segment care more about service attributes than the travellers in other segments. The most favoured alternative in this segment was bus carrier D (as indicated by the carrier-specific constant). 20% of the bus travellers were in Segment 2. Bus travellers in this segment generally preferred carrier A. The coefficients of travel time, frequency and on-time performance were more sensitive in this segment than in other segments. Bus travellers in Segment 3 (24%) were very sensitive to bus fare and insensitive to other attributes except for on-time performance and service attitude. Segment 4 consisted of 23% of bus travellers who were sensitive to bus fare as well as other attributes and generally preferred to use carrier C.

Table 8 reports the estimation result of the LCGNL model with individual characteristics in segment membership functions. The membership coefficients in Segment 4 were set to zero. The results indicate that positive coefficients were related to

Table 7. The LCGNL model result without individual characteristics in membership functions: bus choice.

Variables	Segment 1 Coeff. (*t*-value)	Segment 2 Coeff. (*t*-value)	Segment 3 Coeff. (*t*-value)	Segment 4 Coeff. (*t*-value)
Bus A constant	–0.256 (–3.87)	1.008 (7.28)	0.632 (5.53)	1.946 (5.67)
Bus B constant	–0.227 (–2.84)	0.924 (6.03)	0.861 (7.29)	2.185 (6.29)
Bus C constant	–0.003 (–0.04)	0.803 (5.09)	–0.039 (–0.28)	3.562 (10.15)
Fare	–0.066 (–0.74)	–0.739 (–5.91)	–1.746 (–14.72)	–0.955 (–10.64)
Travel time	–0.900 (–8.14)	–3.302 (–13.79)	–0.181 (–1.36)	–0.516 (–3.44)
Frequency	0.020 (0.76)	0.496 (10.81)	0.014 (0.45)	0.012 (0.35)
Punctuality	0.306 (9.06)	0.459 (8.09)	0.095 (2.27)	0.128 (2.33)
Comfort (comfortable enough)	1.322 (17.74)	0.196 (1.88)	0.075 (0.88)	0.347 (2.95)
Comfort (very comfortable)	1.824 (25.35)	1.084 (9.37)	0.130 (1.60)	0.551 (5.48)
Driver (safe enough)	1.224 (16.04)	–0.076 (–0.69)	0.100 (1.13)	0.078 (0.67)
Driver (very safe)	1.770 (24.04)	0.938 (8.01)	0.109 (1.42)	0.612 (5.87)
Attitude (friendly enough)	0.725 (9.11)	0.335 (2.91)	0.379 (4.98)	0.239 (2.42)
Attitude (very friendly)	0.903 (12.74)	0.977 (9.56)	0.467 (6.35)	0.606 (5.28)
Logsum (*t*-value vs. 1)				
Bus A and B nest	0.369 (8.85)	0.369 (8.85)	0.369 (8.85)	0.369 (8.85)
Bus A and C nest		0.749 (2.88)	0.749 (2.88)	
Allocation				
Bus A and B nest				
Bus A		0.435 (7.64)	0.435 (7.64)	
Bus B		1.000 (–)	1.000 (–)	
Bus A and C nest				
Bus A		0.565 (9.93)	0.565 (9.93)	
Bus C		1.000 (–)	1.000 (–)	
Bus D		1.000 (–)	1.000 (–)	
Membership function constant	0.366 (2.94)	−0.120 (−0.81)	0.058 (0.41)	
Segment size	33%	20%	24%	23%
Final log-likelihood	–6996.173			
Likelihood ratio	0.2179			
Adjusted likelihood ratio	0.2115			

high-income travellers in Segments 1 and 3. Likewise, bus travellers in Segment 4 were relatively lower income and more likely to choose carriers B and C. The LCGNL model with socioeconomic and trip variables in the segment memberships had the best goodness-of-fit of all the LC models in terms of the log-likelihood, likelihood ratio index and adjusted likelihood ratio index. This LCGNL model outperformed the corresponding LCGNL model without socioeconomic and trip variables in the segment memberships, using the likelihood ratio test at the 0.05 level ($\chi^2 = 13.20 > 7.81$). Thus, the LCGNL model with segment membership functions, including the personal income variables, is the most preferred LC specification.

3.4. *Discussion*

Modelling passenger carrier choice typically used the MNL or NL model in the current literature (Proussaloglou and Koppelman 1995, Hensher *et al*. 2003, Eboli and Mazzulla 2008, Wen *et al*. 2009). Our results indicate that advanced GEV models such as GNL

Table 8. The LCGNL model result with individual characteristics in membership functions: bus choice.

Variables	Segment 1 Coeff. (*t*-value)	Segment 2 Coeff. (*t*-value)	Segment 3 Coeff. (*t*-value)	Segment 4 Coeff. (*t*-value)
Bus A constant	–0.261 (–3.88)	1.007 (7.18)	0.631 (5.47)	2.398 (6.56)
Bus B constant	–0.229 (–2.84)	0.931 (6.01)	0.871 (7.35)	2.611 (7.05)
Bus C constant	0.002 (0.04)	0.808 (5.06)	–0.053 (–0.37)	3.953 (10.56)
Fare	–0.039 (–0.43)	–0.725 (–5.96)	–1.751 (–14.47)	–0.917 (–10.44)
Travel time	–0.904 (–8.14)	–3.269 (–13.74)	–0.188 (–1.38)	–0.504 (–3.44)
Frequency	0.020 (0.74)	0.485 (10.84)	0.016 (0.52)	0.015 (0.43)
Punctuality	0.304 (8.95)	0.461 (8.10)	0.093 (2.21)	0.131 (2.41)
Comfort (comfortable enough)	1.325 (17.73)	0.200 (1.94)	0.070 (0.87)	0.355 (3.07)
Comfort (very comfortable)	1.825 (25.12)	1.088 (9.43)	0.128 (1.52)	0.559 (5.69)
Driver (safe enough)	1.231 (15.99)	–0.077 (–0.70)	0.113 (1.25)	0.078 (0.69)
Driver (very safe)	1.780 (23.86)	0.947 (8.09)	0.105 (1.36)	0.608 (5.95)
Attitude (friendly enough)	0.720 (9.02)	0.341 (3.00)	0.396 (5.12)	0.241 (2.51)
Attitude (very friendly)	0.897 (12.59)	0.983 (9.61)	0.477 (6.40)	0.602 (5.30)
Logsum (*t*-value vs. 1)				
Bus A and B nest	0.362 (9.11)	0.362 (9.11)	0.362 (9.11)	0.362 (9.11)
Bus A and C nest		0.765 (2.62)	0.765 (2.62)	
Allocation				
Bus A and B nest				
Bus A		0.433 (8.54)	0.433 (8.54)	
Bus B		1.000 (–)	1.000 (–)	
Bus A and C nest				
Bus A		0.567 (10.20)	0.567 (10.20)	
Bus C		1.000 (–)	1.000 (–)	
Bus D		1.000 (–)	1.000 (–)	
Membership function				
Constant	0.115 (0.78)	−0.170 (−1.02)	−0.210 (−1.23)	
Income	0.640 (2.52)	0.146 (0.50)	0.661 (2.40)	
Segment size	33%	20%	24%	23%
Final log-likelihood	−6989.58			
Likelihood ratio	0.2187			
Adjusted likelihood ratio	0.2119			

appear to be superior to the standard MNL and NL in modelling carrier choice behaviour. For example, two Japanese airlines (ANA and JAL) may have some similar characteristics. On the other hand, airlines that perform strategic cooperation under a code-sharing agreement, such as ANA and EVA, may be perceived by air travellers to be similar in certain features. The complex substitution patterns among airlines were evidently elucidated by the GNL model.

The proposed LCGNL model allows a flexible structure for the covariance among alternatives as well as preference heterogeneity across individuals, while simultaneously identifying segment sizes and profiles. In both empirical cases, the LCGNL model statistically outperformed the corresponding LCMNL and LCNL models, illustrating the applicability of the proposed model. Besides, the LCGNL model with individual characteristics in segment membership functions had the best goodness-of-fit of all the LC models; therefore, it is a better modelling approach for analysing carrier choice behaviour.

In both cases, the estimations of the LCGNL models imposed equality constraints on the logsum and allocation parameters across segments, which results in decreasing the significance of some logsum estimates, especially in the case of a large number of segments (e.g. bus choice model). This phenomenon often occurs when accounting for the joint effects of taste variation and flexible substitution patterns between alternatives (Hess *et al.* 2005). Such a problem may be resolved by allowing differential logsum and allocation parameters across segments and imposing restrictions on these estimates within a reasonable range. Consequently, the result gives some segments a GNL structure and the others an NL structure in which the model still fits into an LCGNL structure.

The estimation results of various models indicate that both air and bus travellers perceived service quality of carriers to be important. This is consistent with the previous findings reported in Eboli and Mazzulla (2008), Espino *et al.* (2008), Martín *et al.* (2008) and Habib *et al.* (2011). Accordingly, carriers can use the findings to develop effective operational and marketing strategies. The check-in service for air travel, for example, is considered to be important based on the relatively large coefficient estimates. The interpretation is that airline check-in is the earliest service that air passengers encounter when arriving at the airport. For international travel, passengers are required to check-in at the counter 2 hours before departure. Long queues and waiting time are very likely to occur during peak hours at the airport, resulting in poor service quality perceived by passengers. Taoyuan International Airport Authority and some airlines have initiated major improvements of check-in service at the airport. Since December 2009, two Taiwanese airlines (CAL and EVA) allow passengers to receive their boarding pass and pre-check their luggage at a Taiwan high-speed rail station near the international airport. Additionally, a new construction of Taoyuan International Airport Metro line is underway. Similar to the Hong Kong International Airport, self-service check-in kiosks will be installed at some Metro stations to increase travel convenience. Airlines could also popularise online check-in service allowing prior check-in within 24 hours of flight departure. Although kiosks can reduce check-in times and queues at check-in counters, most interfaces are not friendly enough for air travellers, particularly first-time users. Airline companies should allocate staff, especially during peak hours, to assist air travellers with self-service check-in kiosks.

This study has identified separate segments for air and bus markets. Strategies are likely to be effective when accounting for the travellers' preferences of each segment. For example, air travellers in Segment 2 are sensitive to price, although they are also concerned about flight frequency. This group of air travellers considers airfare to be a critical factor when choosing airlines. Differential and discount airfares may be effective for customers in this segment. All international airlines in this route are full-service operations, and low-cost airlines do not exist. To attract price-sensitive but quality-insensitive air travellers, airlines could launch low-cost services.

4. Conclusions

This article proposes a new LC model with the GNL formulation – the LCGNL model – which performs two functions: it accommodates flexible substitution patterns among alternatives and heterogeneous preferences for individuals, and it identifies segment sizes and individual profiles. The LCGNL model extends the popular LCMNL model, such that the IIA property does not hold within segments. The LCNL model, an extension of the LCMNL model, allows the correlated errors of utilities for alternatives in a common nest but still suffers from the disadvantages of the NL model. By integrating the GNL, one of the most flexible GEV models, into the LC model, the LCGNL shows considerably higher flexibility in error covariance structure than the LCNL model. The development of the LCGNL model has made a methodological contribution to the literature.

The LCGNL and other models were empirically demonstrated using airline and bus choice datasets. Both empirical cases indicate that the conventional MNL or NL models may fail to capture complex substitution patterns among carriers. Likewise, the GNL model provides a more flexible covariance structure for modelling competition among carriers. The estimation results of various LC models indicate the existence of taste heterogeneity in both cases. The LC model adequately accommodates variations of taste parameters as well as explicitly identifies the number, sizes and characteristics of segments.

The proposed LCGNL model statistically outperformed the LCMNL and LCNL models in both cases, which demonstrates the applicability and superiority of this model for analysing carrier choice behaviour. The LCGNL model, with segment membership functions that incorporated individual socioeconomic and trip characteristics, can further enhance the goodness-of-fit and is the most preferred specification. Based on the estimation result of the preferred model, carriers can develop effective operational and marketing strategies to attract new customers.

A number of ideas could be considered for future research. This study chooses two data sets for empirical illustrations. More applications in other geographic regions of interest with large data sets to validate the proposed LCGNL model would be helpful. This research has used various LC models to analyse carrier market segmentation. Once customers' segments have been identified, carriers should select target market(s) and understand competitive positions relative to their rivals. Positioning analysis often uses a choice mapping approach to illustrate the competitive positions of products or services (Chintagunta 1994, González-Benito *et al*. 2009, Yang and Sung 2010). Future studies could integrate the LCGNL model with the choice mapping approach to simultaneously explore market segmentation and competitive positioning within the carrier choice context.

Acknowledgements

The authors thank Professor S.C. Wong (Editor-in-Chief) and the reviewers for their insightful suggestions and comments. This study was financially supported by the National Science Council (NSC) of Taiwan under Grant No. NSC 96-2416-H-035-005-MY2.

References

Aptech Systems, 2008. *GAUSS version 9.0 user guide*. Black Diamond, WA: Aptech Systems.

Arunotayanun, K. and Polak, J.W., 2011. Taste heterogeneity and market segmentation in freight shippers' mode choice behaviour. *Transportation Research Part E*, 47 (2), 138–148.

Badoe, D.A. and Miller, E.J., 1998. An automated segmentation procedure for studying variations in mode choice behavior. *Journal of Advanced Transportation*, 32 (2), 190–215.

Bajwa, S., *et al*., 2008. Discrete choice modeling of combined mode choice and departure time. *Transportmetrica*, 4 (2), 155–177.

Bekhor, S. and Elgar, A., 2007. Investment in mobility by car as an explanatory variable for market segmentation. *Journal of Public Transportation*, 10 (2), 17–32.

Bekhor, S. and Prashker, J.N., 2008. GEV-based destination choice models that account for unobserved similarities among alternatives. *Transportation Research Part B*, 42 (3), 243–262.
Ben-Akiva, M. and Bierlaire, M., 1999. Discrete choice methods and their applications to short-term travel decisions. *In*: R. Hall, ed. *Handbook of transportation science*. Dordrecht, The Netherlands: Kluwer Academic, 5–34.
Bhat, C.R., 1997. An endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science*, 31 (1), 34–48.
Bhat, C.R., 1998. An analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B*, 32 (6), 361–371.
Bhat, C.R. and Guo, J., 2004. A mixed spatially correlated logit model: formulation and application to residential choice modeling. *Transportation Research Part B*, 38 (2), 147–168.
Bierlaire, M., 2002. The network GEV model. *In*: *Proceedings of the 2nd Swiss transportation research conference*, Ascona, Switzerland.
Bierlaire, M., 2006. A theoretical analysis of the cross-nested logit model. *Annals of Operations Research*, 144 (1), 287–300.
Bodapati, A.V. and Gupta, S., 2004. The recoverability of segmentation structure from store-level aggregate data. *Journal of Marketing Research*, 41 (3), 351–364.
Bucklin, R.E. and Gupta, S., 1992. Brand choice, purchase incidence, and segmentation: an integrated modeling approach. *Journal of Marketing Research*, 29 (2), 201–215.
Chintagunta, P.K., 1994. Heterogeneous logit model implications for brand positioning. *Journal of Marketing Research*, 31 (2), 304–311.
Chu, C., (1989). A paired combinatorial logit model for travel demand analysis. *In*: *Proceedings of the fifth world conference on transportation research*, Ventura, CA, Vol. 4, 295–309.
Daly, A. and Bierlaire, M., 2006. A general and operational representation of generalised extreme value models. *Transportation Research Part B*, 40 (4), 285–305.
Depaire, B., Wets, G., and Vanhoof, K., 2008. Traffic accident segmentation by means of latent class clustering. *Accident Analysis and Prevention*, 40 (4), 1257–1266.
Eboli, L. and Mazzulla, G., 2008. A stated preference experiment for measuring service quality in public transport. *Transportation Planning and Technology*, 31 (5), 509–523.
Espino, R., Martín, J.C., and Román, C., 2008. Analyzing the effect of preference heterogeneity on willingness to pay for improving service quality in an airline choice context. *Transportation Research Part E*, 44 (4), 593–606.
González-Benito, O., Martínez-Ruiz, M.P., and Mollá-Descals, A., 2009. Using store level scanner data to improve category management decisions: developing positioning maps. *European Journal of Operational Research*, 198 (2), 666–674.
Greene, W.H. and Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B*, 37 (8), 681–698.
Gupta, S. and Chintagunta, K., 1994. On using demographic variables to determine segment membership in logit mixture models. *Journal of Marketing Research*, 31 (1), 128–136.
Habib, K.M.N., Kattan, L., and Islam, M.T., 2011. Model of personal attitudes towards transit service quality. *Journal of Advanced Transportation*, 45 (4), 271–285.
Hensher, D.A., Stopher, P., and Bullock, P., 2003. Service quality-developing a service quality index in the provision of commercial bus contracts. *Transportation Research Part A*, 37 (6), 499–517.
Hess, S., Bierlaire, M., and Polak, J.W., 2005. Capturing correlation and taste heterogeneity with mixed GEV models. *In*: R. Scarpa and A. Alberini, eds. *Applications of simulation methods in environmental and resource economics*. Dordrecht, The Netherlands: Springer, 55–76.
Kamakura, W.A. and Russell, G.A., 1989. A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, 26 (4), 379–390.
Kamakura, W.A., Kim, B.-D., and Lee, J., 1996. Modeling preference and structural heterogeneity in consumer choice. *Marketing Science*, 15 (2), 152–172.

Koppelman, F.S. and Wen, C.-H., 2000. The paired combinatorial logit model: properties, estimation and application. *Transportation Research Part B*, 34 (2), 75–89.

Loo, B.P.Y., 2008. Passengers' airport choice within multi-airport regions (MARs): some insights from a stated preference survey at Hong Kong International Airport. *Journal of Transport Geography*, 16 (2), 117–125.

Loo, B.P.Y., Wong, S.C., and Hau, T.D., 2006. Introducing alternative fuel vehicles in Hong Kong: views from the public light bus industry. *Transportation*, 33 (6), 605–619.

Martín, J.C., Román, C., and Espino, R., 2008. Willingness to pay for airline service quality. *Transport Reviews*, 28 (2), 199–217.

McFadden, D., 1978. Modeling the choice of residential location. *Transportation Research Record*, 672, 72–77.

McFadden, D. and Train, K., 2000. Mixed MNL models of discrete response. *Journal of Applied Econometrics*, 15 (5), 447–470.

Newman, J.P., 2008. Normalization of network generalized extreme value models. *Transportation Research Part B*, 42 (10), 958–969.

Outwater, M.L., *et al.*, 2004. Attitudinal market segmentation approach to mode choice and ridership forecasting. *Transportation Research Board*, 1854, 32–42.

Papola, A., 2004. Some developments on the cross-nested logit model. *Transportation Research Part B*, 38 (9), 833–851.

Pels, E., Nijkamp, P., and Rietveld, P., 2001. Airport and airline choice in a multiple airport region: an empirical analysis for the San Francisco bay Area. *Regional Studies*, 35 (1), 1–9.

Proussaloglou, K. and Koppelman, F.S., 1995. Air carrier demand. *Transportation*, 22 (4), 371–388.

Psaraki, V. and Abacoumkin, C., 2002. Access mode choice for relocated airports: the new Athens International Airport. *Journal of Air Transport Management*, 8 (2), 89–98.

Rastogi, R. and Krishna Rao, K.V., 2009. Segmentation analysis of commuters accessing transit: Mumbai study. *Journal of Transportation Engineering*, 135 (8), 506–515.

Revelt, D. and Train, K., 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*, 80 (4), 648–657.

Small, K., 1987. A discrete choice model for ordered alternatives. *Econometrica*, 55 (2), 409–424.

Swait, J., 2003. Flexible covariance structures for categorical dependent variables through finite mixtures of generalized extreme value models. *Journal of Business and Economic Statistics*, 21 (1), 80–87.

Teichert, T., Shehu, E., and von Wartburg, I., 2008. Customer segmentation revisited: the case of airline industry. *Transportation Research Part A*, 42 (1), 227–242.

Vovsha, P., 1997. The cross-nested logit model: application to mode choice in the Tel-Aviv metropolitan Area. *Transportation Research Record*, 1607, 6–15.

Walker, J.L. and Li, J., 2007. Latent lifestyle preferences and household location decisions. *Journal of Geographical Systems*, 9 (1), 77–101.

Wen, C.-H., Chen, T.-N., and Huang, W.-W., 2009. Mixed logit analysis of international airline choice: an empirical study of Taiwan's air routes. *Transportation Research Record*, 2106, 20–29.

Wen, C.-H. and Koppelman, F.S., 2001. The generalized nested logit model. *Transportation Research B*, 35 (7), 627–641.

Wen, C.-H. and Lai, S.-C., 2010. Latent class models of international air carrier choice. *Transportation Research Part E*, 46 (2), 211–221.

Wen, C.-H., *et al.*, 2011. A latent class nested logit model of intercity bus carrier choice. *Journal of the Eastern Asia Society for Transportation Studies*, 9, 382–394.

Yang, C.-W. and Sung, Y.-C., 2010. Constructing a mixed-logit model with market positioning to analyze the effects of new mode introduction. *Journal of Transport Geography*, 18 (1), 175–182.

Zhang, J., *et al.*, 2009. Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms. *Transportation Research Part B*, 43 (2), 230–250.