Pergamon

Transpn Res.-B, Vol. 31, No. 1, pp. 11-21, 1997

0191-2615/96 $17.00 + 0.00

PII:S0191-2615(96)00018-5

COVARIANCE HETEROGENEITY IN NESTED LOGIT MODELS: ECONOMETRIC STRUCTURE AND APPLICATION TO INTERCITY TRAVEL

CHANDRA R. BHAT

Department of Civil and Environmental Engineering, University of Massachusetts, Amherst, MA 01003, U.S.A.

(*Received 16 October 1995; in revised form 12 April 1996*)

Abstract—The nested logit model relaxes the 'independence of irrelevant alternatives' (IIA) property of the multinomial logit model by grouping alternatives based on their degree of substitution. Alternatives in a nest exhibit an identical degree of increased sensitivity relative to alternatives not in the nest. An assumption maintained in the nested logit is that the degree of sensitivity among nested alternatives is invariant across agents making the choice. This assumption might be untenable in many situations. In this paper, we propose an extension of the nested logit model to allow heterogeneity (across agents) in the covariance among nested alternatives based on observed agent characteristics. We label this model as the COVNL model. The multinomial logit, the nested logit, and the COVNL model are estimated to examine the impact of improved rail service on weekday, business travel in the Toronto–Montreal corridor. The empirical results show that not accounting for covariance heterogeneity in the nested logit formulation leads to a statistically inferior data fit and also to biased model estimates of the effects of level-of-service variables. Copyright © 1996 Elsevier Science Ltd

1. INTRODUCTION

Discrete choice models are used to predict the choice of an alternative from among many discrete alternatives based on the characteristics of the agent making the choice and the attributes of each available alternative. Discrete choice models are now used very commonly, and with increasing sophistication, in the analysis of such diverse choices (to list a few) as purchase incidence and brand choice in marketing (see Jain *et al.*, 1994; Bucklin *et al.*, 1995, etc.), travel mode and car ownership choice in transportation (see Bhat, 1995; Hensher *et al.*, 1992, etc.), housing type and location choice in geography (Fischer & Aufhauser, 1988; Evers, 1990), labor force participation and labor migration in labor economics (Ashenfelter & Layard, 1986; Nijkamp, 1986), and investment choices of firms in finance (Corres *et al.*, 1993).

In most discrete choice applications, McFadden's (1973) multinomial logit formulation is used. The multinomial logit (MNL) model has a simple and elegant closed-form mathematical structure, making it easy to estimate and interpret. However, it is saddled with the 'independence of irrelevant alternatives' (IIA) property at the individual level; that is, the multinomial logit model imposes the restriction of equal cross-elasticities due to a change in an attribute affecting only the utility of an alternative i for all alternatives $j \neq i$. This property of equal proportionate change of unchanged modes is unlikely to represent actual choice behavior in many situations (Stopher *et al.*, 1981).

One approach to relaxing the restrictive IIA property of the multinomial logit is to allow a completely free correlation structure (subject to some identifiability conditions) among the random error components of utility of each discrete choice alternative. This results in the multinomial probit (MNP) model of choice. Unfortunately, the increase in flexibility of error structure in the multinomial probit comes at the expense of introducing several additional parameters in the covariance matrix. This generates a number of conceptual, statistical and practical problems, including difficulty in interpretation, highly nonintuitive model behavior, low precision of covariance parameter estimates, and increased difficulty in transferring models from one space-time sampling frame to another (see Horowitz, 1991; Currim, 1982). The multinominal probit choice probabilities also involve high dimensional integrals and this poses computational problems when the number of alternatives exceeds four.*

A compromise between the multinomial logit model and multinomial probit model in terms of flexibility and computational/interpretational ease is the nested logit model (Daly & Zachary, 1979; McFadden, 1978). This model permits covariance in random components among subsets of alternatives and, thus, is free of the restrictive IIA property of the MNL model. Also, though the estimation of this model is not as straightforward as that of the MNL, it does have a closed form solution for the choice probabilities and is relatively simple to estimate. The nested logit (NL) model is seeing increasing use in recent years in multidimensional choice contexts (for example, see Waddell 1993; Evers, 1990) and in unidimensional contexts where subsets of the available alternatives share common unobserved components of utility (for example, see Forinash & Koppelman, 1993; Brownstone & Small, 1989). While earlier applications of the nested logit used a two-stage sequential procedure for estimation, the more recent applications use an efficient full-information maximum likelihood procedure.

The nested logit formulation used in all earlier applications, to the author's knowledge, imposes the restriction of equal correlation in the random utility components among nested alternatives across agents. This restriction of equal correlation across agents might be untenable in many choice contexts and, at the least, should be tested. For example, consider an intercity travel mode choice model with car and train grouped as surface modes and air treated as a non-nested alternative. The degree of (increased) sensitivity (or cross-elasticity) between the two surface transportation modes relative to the air mode may differ based on characteristics of the traveler such as income (lower income may imply greater sensitivity between the surface modes) and attributes of the traveler's trip such as trip distance (shorter trip distances may lead to greater sensitivity between surface modes). In general, ignoring the variation in the covariance among nested alternatives across agents making the choice (i.e. ignoring covariance heterogeneity in a nested logit model) will lead to biased and inconsistent estimates of the effects of relevant policy-sensitive and other variables (covariance heterogeneity is one particular form of population heterogeneity; ignoring population heterogeneity of any form when it is present leads to biased and inconsistent estimation; see Chamberlain, 1980; Revelt & Train, 1995).

In this paper, we formulate and estimate a nested logit model which accommodates covariance heterogeneity (COVNL model). The author is unaware of any previous research in the discrete choice field that formulates and estimates such a model. The COVNL model retains a simple form and provides closed-form expressions for the choice probabilities. The performance of the COVNL model is empirically compared with that of the standard nested logit (NL) model and the multinomial logit model in the context of intercity travel mode choice.

The rest of this paper is structured as follows. The next section develops the structure and presents the estimation procedure for the COVNL model. Section 3 discusses the data and the empirical results. The final section provides a summary of the research findings.

2. MODEL STRUCTURE AND ESTIMATION

We develop the model structure in the context of a simple two-level, three alternative, nested structure to simplify the notation. In this structure, we assume that alternatives b and c are nested together and that alternative a is the non-nested alternative. Also, to simplify the presentation, we assume that all agents have all alternatives available to them. Extensions of the formulation to more complicated two-level nested structures and

*However, these computational barriers have been significantly reduced by recent developments that use a smooth recursive conditioning (SRC) simulator (Börsch-Supan & Hajivassiliou, 1993).

to the case where some individuals have only a subset of alternatives available to them is straightforward. Extensions to higher level nested structures is discussed later in the section.

Following the usual random utility framework for discrete choice models, we express the utility U_{qi} that an individual q (we will use the terms individual and agent interchangeably in this paper) associates with an alternative i as the sum of a deterministic component V_{qi} (that depends on observed attributes of the alternative and the individual) and a random component ϵ_{qi} (that represents the effects of unobserved attributes of the individual and unobserved characteristics of the alternative). We assume that the ϵ_{qi} values have a type I extreme value distribution and are identically distributed (across alternatives and across individuals) with a location parameter equal to zero and a scale parameter equal to 1.* The ϵ_{qi} values for the nested alternatives (i = b,c) are decomposed into a common (across the nested alternatives) error component, ϵ_{qg}, and an independent (across the nested alternatives) error component, ϵ_{qgi}. The common and independent error components are independently distributed across individuals and are independent of each other for all q. The independent error components are assumed to be identically distributed across alternatives b and c with a type I extreme value distribution and with a location parameter equal to zero and a scale parameter equal to θ_q. In equation form,

$$\begin{aligned} U_{qa} &= V_{qa} + \epsilon_{qa}, && \epsilon_{qa} \sim \lambda(0,1) \text{ for all } q \\ U_{qb} &= V_{qb} + \epsilon_{qg} + \epsilon_{qgb}, && (\epsilon_{qg} + \epsilon_{qgb}) \sim \lambda(0,1) \text{ for all } q,\ \epsilon_{qgb} \sim \lambda(0,\theta_q) \\ U_{qc} &= V_{qc} + \epsilon_{qg} + \epsilon_{qgc}, && (\epsilon_{qg} + \epsilon_{qgc}) \sim \lambda(0,1) \text{ for all } q,\ \epsilon_{qgc} \sim \lambda(0,\theta_q) \end{aligned} \tag{1}$$

where $\lambda(\mu,\omega)$ represents the probability density function of the type I extreme value distribution with location parameter μ and scale parameter ω. The presence of the common error component ϵ_{qg} in the random utilities of alternatives b and c generates a covariance:

$$\mathrm{Cov}(U_{qb},U_{qc}) = \mathrm{Var}(\epsilon_{qg}) = \frac{\pi^2}{6}[1 - \theta_q^2], \text{ or equivalently } \mathrm{Cor}(U_{qb},U_{qc}) = 1 - \theta_q^2. \tag{2}$$

Since each (and both of) the scale parameter θ_q and the variance of ϵ_{qg} cannot be negative, the condition $0<\theta_q \leq 1$ for all q is a necessary and sufficient condition for the COVNL model to be interpreted as a random utility model over the entire domain of the I-variate real-space (R^I) characterized by the utilities of the I alternatives (see McFadden, 1978).†

The nested-logit equation system in eqn (1) yields the following expressions for the choice probabilities:

$$P_{qa} = \frac{\exp(V_{qa})}{\exp(V_{qa}) + \exp(\theta_q \ln\Gamma_q)},\ \Gamma_q = [\exp(V_{qb}/\theta_q) + \exp(V_{qc}/\theta_q)]$$

$$P_{qj} = \frac{\exp(\theta_q \ln\Gamma_q)}{\exp(V_{qa}) + \exp(\theta_q \ln\Gamma_q)} \cdot \frac{\exp(V_{qj}/\theta_q)}{\Gamma_q},\ j = \text{b,c}. \tag{3}$$

We assume a linear-in-parameters specification for the systematic utility of each alternative given by $V_{qi} = \beta' X_{qi}$ for the qth individual and ith alternative (i = a,b,c). In eqn (3), if we

*We maintain the assumption of identical variance of the random utilities of the different alternatives in this paper, a condition required in the theoretical derivation of the nested logit model (see McFadden, 1978). Interestingly, Börsch-Supan (1990) has observed recently that, from an empirical standpoint, it appears more important to group alternatives in the nested logit model according to differences among variances of the random utilities than inter-alternative correlation. Börsch-Supan's result, however, is based on a rather small simulated sample of 100 observations; further theoretical and empirical research into Börsch-Supan's preliminary findings is needed.

†Börsch-Supan (1990) has recently shown for the case of a standard nested logit model (where $\theta_q = \theta$ for all q) that the condition $\theta \leq 1$ is not necessary if the domain of all potential deterministic utilities fall within a certain subset of R^I (this subset being determined by θ). He presents the weaker, data-specific, validity conditions under this restrictive situation. In this paper, we maintain the condition $0<\theta_q \leq 1$ for all q which guarantees consistency of the nested model with random utility maximization in the full domain of R^I and which is independent of the data context.

constrain θ_q to be equal across individuals, we obtain the standard nested logit model. As indicated earlier, such a constraint may be restrictive since the sensitivity between the nested alternatives (equivalently, the correlation between the utilities of the alternatives b and c) may be a function of relevant observed individual characteristics. To allow for such covariance heterogeneity, we relate θ_q to observed individual characteristics as follows:

$$\theta_q = F(\alpha + \gamma'\mathbf{z}_q) \tag{4}$$

where $\mathbf{z}_q$ is a vector of individual characteristics, γ is a corresponding vector of parameters to be estimated, and F is a transformation function. The argument of F in eqn (4) can in principle take any value on the real line, but θ_q must be in the 0–1 range. Further, an increase in any element of the vector $\mathbf{z}_q$ should have a monotonic effect on θ_q. The key is to use a transformation function F that has the properties:

$$F(-\infty) = 0,\ F(+\infty) = 1, \text{ and } f(x) = \frac{\partial F(x)}{\partial x} > 0. \tag{5}$$

Thus, F is a continuous, monotonically increasing function that maps from the real line to the 0–1 interval. Any proper continuous cumulative probability distribution function will satisfy these properties.

If $\gamma = 0$ in eqn (4), then it implies the absence of covariance heterogeneity and the COVNL model collapses to the standard nested logit model. Thus, the COVNL model can be tested against the standard nested logit model using usual nested likelihood ratio tests.*

A useful point to note regarding the proposed COVNL model is that the covariance heterogeneity is introduced by adopting a heteroscedastic specification (across individuals) for the random components, ϵ_{qgb} and ϵ_{qgc}, in the conditional choice model (i.e. in the model of choice of nested alternative given that one of the nested alternatives is chosen). Such a heteroscedastic specification is appropriate when the uncertainty in an analyst's ability to ascertain the systematic component of utility of the nested modes, given that one of the nested modes is chosen, varies depending upon observed characteristics of individuals. For example, in an intercity mode choice context where ground modes (say train and car) are nested together against the air mode, the perception of high income ground-mode users regarding comfort and convenience of the ground modes may vary substantially more than the corresponding perception of low income ground-mode users. This might be because the air mode is less accessible as an alternative mode to low income ground-mode users than high income ground-mode users. Consequently, the low income individuals may attune their perceptions of comfort and convenience of the ground modes so as to be more uniformly satisfactory than the distribution of perceptions of high income individuals. Since individual perceptions of comfort and convenience of modes are unobserved attributes, this would lead to heteroscedasticity across individuals in the random error components of the mode utilities in the conditional (on ground mode) choice model.

The nested logit model with covariance heterogeneity is estimated using the maximum likelihood technique. The parameters to be estimated include the β vector in the systematic utility component, and the alpha and γ parameters determining the covariance among the nested alternatives. The log-likelihood function to be maximized takes the familiar form:

$$\mathcal{L} = \sum_q \sum_i y_{qi} \log P_{qi}, \text{ where} \tag{6}$$

$$y_{qi} = \begin{cases} 1 \text{ if the } q\text{th individual chooses alternative } i \\ 0 \text{ otherwise,} \end{cases} \qquad (q = 1, 2, \ldots Q,\ i = \text{a,b,c}).$$

*A theoretical advantage of the transformation in eqn (4) even in the context of the standard nested logit model (i.e. when γ is restricted to zero) is that it implicitly guarantees the range of the logsum parameter to be between 0 and 1. From a practical viewpoint, however, we found that whenever the logsum parameter exceeded one in the untransformed NL specification, the estimated α was a large positive value and $F(\alpha)=1$ for all practical purposes in the transformed specification (thus, collapsing the specification to the multinomial logit).

Prior to maximizing the likelihood function above, the distribution function F in eqn (4) needs to be specified. Any continuous probability distribution function may be used; the specific functional form chosen is not very critical. In this research, we employ the logistic distribution because it has a simple form and also because it approaches the boundary values of 0 and 1 more slowly relative to other distributions such as the normal (the latter property of the logistic distribution makes θ_q sensitive to the linear index $\alpha + \gamma'\mathbf{z}_q$ over a wider range of the real line and appeared to improve the convergence characteristics of the log-likelihood function for some nested logit specifications).

Maximization of the log-likelihood function was achieved using the GAUSS matrix programming language. The analytical gradients of the log-likelihood function with respect to the parameters were coded.

The above formulation and estimation has focused on a two-level nested structure. However, one can include covariance heterogeneity in higher level nested structures too. For example, in a three-level nested structure, there are two logsum parameters; the lower nest logsum parameter θ_{ql}, which determines the covariance among alternatives in the lowest nest and an upper nest logsum parameter θ_{qu}, which determines the covariance among alternatives in the upper nest. The logsum parameters should satisfy the condition $1 \geq \theta_{qu} \geq \theta_{ql} > 0$ (for each and all q) to be compatible with utility maximization in the entire domain of R^I.* The inequality conditions on the logsum parameters can be maintained by using the following specifications:

$$\theta_{ql} = F[\alpha + \gamma'\mathbf{z}_q] \text{ and } \theta_{qu} = F[(\alpha + \gamma'\mathbf{z}_q) + G(\delta + \eta'\mathbf{w}_q)] \tag{7}$$

where $\mathbf{w}_q$ is a vector of individual characteristics ($\mathbf{w}_q$ can be the same as $\mathbf{z}_q$), δ and η are additional parameters to be estimated in the three-level nested structure, F is a continuous cumulative distribution function as earlier, and G is a continuous, monotonically increasing, transformation function that maps from the real line to the 0–∞ interval (thus, $G(-\infty) = 0$ and $G(+\infty) = \infty$; a simple functional form for G which satisfies these properties is the exponential form). The term $G(\delta + \eta'\mathbf{w}_q)$ in eqn (7) determines the differential correlation among alternatives in the upper and lower nests for individual q. As this term approaches zero, that is as $(\delta + \eta'\mathbf{w}_q) \to -\infty$, $\theta_{qu} \to \theta_{ql}$ and the upper nest and lower nest alternatives can be grouped in a single nest. If $(\alpha + \gamma'\mathbf{z}_q) \to +\infty$, the individual's choice structure is represented by the multinomial logit.

3. AN APPLICATION TO INTERCITY MODE CHOICE MODELING

3.1. Background and model specification

Substantial attention has been focused in recent years on identifying and evaluating alternative proposals to improve intercity transportation services in response to increasing congestion in intercity travel, and the consequent adverse impacts of such congestion on economic development and environmental quality (for example, see Moon, 1991; KPMG Peat Marwick *et al.*, 1993). A critical issue in the evaluation and effective implementation of these proposals is the development of mode choice models which are able to capture the competitive structure among modes.

In this section, we report the empirical results obtained from applying the multinomial logit model, the standard nested logit model, and the nested logit model with covariance heterogeneity (COVNL model) proposed in this paper to an analysis of intercity mode choice behavior for weekday business travel in the Toronto–Montreal corridor of Canada. The travel mode alternatives are air, train, and car. The data used in the current empirical analysis were assembled by VIA Rail (the Canadian national rail carrier) in 1989 (see KPMG Peat Marwick & Koppelman, 1990 for a detailed description of this data). The sample used is choice-based and comprises 2704 business travelers. The aggregate proportion of the choice of each mode in the population of travelers was available from supplementary data. Since the sample is choice-based with known aggregate shares, we

*As discussed earlier for the two-level nested structure, weaker validity conditions may suffice when the domain of all relevant deterministic utilities can be restricted to a subset of R^I.

employed the Weighted Exogenous Sample Maximum Likelihood (WESML) method proposed by Manski and Lerman (1977) in estimation. The asymptotic covariance matrix of parameters is computed as $H^{-1}\Delta H^{-1}$, where H is the hessian and Δ is the cross-product matrix of the gradients (H and Δ are evaluated at the estimated parameter values). This provides consistent standard errors of the parameters (Börsch-Supan, 1987).

We estimated seven different models in the study: a multinomial logit model, three possible nested logit structures with no covariance heterogeneity, and the corresponding three nested logit structures with covariance heterogeneity. The three nested structures are: (a) car and train (slow modes) grouped together in a nest which competes against air, (b) train and air (common carriers) grouped together in a nest which competes against car, and (c) air and car grouped together in a nest which competes against train.

A number of different variable specifications were examined to determine the preferred utility function specification. The final mode choice model specification (i.e. the $\mathbf{x}_{qi}$ vector in $\beta'\mathbf{x}_q$) included alternative specific variables for (a) annual household income, (b) female sex of traveler (introduced as a dummy variable; 1 for females and 0 for males), and (c) travel group size (1 for individuals traveling alone and zero otherwise). The level-of-service variables included a large city indicator which identified whether a trip originated and/or terminated in a large city, and modal level-of-service measures (frequency of service, travel cost, in-vehicle travel time and out-of-vehicle travel time). The large city variable may be viewed as a 'proxy' level-of-service measure capturing mode characteristics associated with large cities. The modal level-of-service measures were introduced in the utility specification as frequency of service, total cost, total travel time, and out-of-vehicle travel time divided by the logarithm of trip distance (this specification was found to be statistically better than other alternative ones such as the simple out-of-vehicle time specification and cost deflated by income).

Four variables associated with individual socio-demographics and trip characteristics were available in the data, and all of these were considered for inclusion as determinants of covariance heterogeneity (i.e. in the $\mathbf{z}_q$ vector in eqn 4). The variables are: annual household income, sex of traveler, travel group size (traveling alone or traveling in a group) and logarithm of (one-way) trip distance. Of these, sex of traveler did not have a statistically significant impact on the covariance between nested alternatives in all the possible nested structures and hence was dropped from the covariance heterogeneity specification.

Table 1 presents the market shares of each mode and the descriptive statistics for the modal level-of-service measures in the sample. The mean statistics for other exogenous variables are: large city indicator (0.698), income (Canadian $54,939.10), female dummy variable (0.177), dummy variable for traveling alone (0.709), and (one-way) trip distance (350.42 km).

3.2. *Empirical results*

In the empirical analysis, we estimated the multinomial logit model, and the standard nested logit and the COVNL model for all the three possible nested structures mentioned earlier. Before estimating the nested structures, we tested the IIA assumption of the multinomial logit model (against the alternative hypothesis of non-IIA) using the chi-squared test proposed by Hausman and McFadden (1984). The test involves comparing parameter estimates obtained from a restricted choice set with those obtained from the full choice set. The test statistic ranged from 97.7 to 124.3 for the three possible restricted binary choice sets. A comparison of the test statistic with the chi-squared distribution

Table 1. Mode market shares and descriptive statistics of modal level-of-service variables

Mode	Share	Frequency (departures/day)	Total cost (in Canadian $)	In-vehicle time (in min.)	Out-of-vehicle time (in min.)
Train	0.10	4.12 (2.2)	56.40 (17.0)	230.85 (107.9)	86.62 (22.2)
Air	0.38	24.58 (13.3)	154.94 (20.3)	55.67 (17.6)	107.31 (25.6)
Car	0.52	not applicable	66.58 (31.3)	237.25 (103.2)	0.00 (0.0)

Notes: The numbers in the table for the level-of-service measures denote one-way mean (no parentheses) and standard deviation (parentheses) values. Air and train frequency refer to total departures by all carriers.

with 9 d.f. (9 coefficients are identifiable in the restricted choice set estimations) very strongly rejects the IIA assumption in all the three restricted choice set tests. Based on this result, we proceeded to consider the nested logit models which relax the IIA assumption. The standard nested logit (NL) model with train and air grouped together obtained a logsum parameter that exceeded one (though not significantly different from one). When this standard nested model was estimated using the transformation $\theta = F(\alpha)$, we obtained an estimate of $\alpha = 13.86$; the implied value of the logsum parameter θ is indistinguishable from one for practical purposes, thus collapsing the structure to the multinomial logit. Inclusion of variables to represent covariance heterogeneity in this nested structure did not improve the log-likelihood function significantly. The NL model with car and air grouped together obtained an acceptable logsum parameter (0.92), but this parameter was not significantly different from one. Incorporation of covariance heterogeneity did not improve the model fit significantly for this structure too. The NL model with train and car grouped together obtained an acceptable logsum parameter (0.78). This parameter was also significantly different from one. Adding covariance heterogeneity to this NL structure led to a further significant improvement in the log-likelihood function.*

The results for the multinomial logit model, the standard nested logit (NL) model with car and train grouped together, and the corresponding nested model with covariance heterogeneity (COVNL model) are shown in Table 2. *Prima facie*, the results of the mode choice parameters are similar in all the three models. In particular, the signs of all the mode choice parameters are consistent with *a priori* expectations (the car mode is used as the base for the alternative specific constants and alternative specific variables). The income parameters show that higher income favors air travel relative to other modes, and low income favors train travel (the latter effect, however, is statistically insignificant). Individuals who travel alone prefer common carrier modes more so than individuals traveling in groups (who prefer the car mode). Women are more likely to use common carrier modes (especially the train mode) compared to men (see KPMG Peat Marwick *et al.*, 1993 for a similar result in the context of a Florida intercity mode choice study). There is a preference for the common carrier modes (air and train) over the car mode, and for the train mode over the air mode, for trips which originate and/or end in a large city. All level-of-service measures yield reasonable parameters. The negative coefficient on the 'out-of-vehicle travel time divided by logarithm of distance' (OVTT/Logdist) variable reveals that out-of-vehicle travel time is more onerous than in-vehicle travel time.

The mode choice parameter estimates from the three models are close to each other. However, there are some significant differences (all significance tests are conducted at the 0.1 level). The COVNL model indicates a lower (and insignificant) preference for train (relative to car) for individuals traveling alone compared to the other two models. The NL and COVNL models suggest a lower magnitude of preference for the common carrier modes (especially for the train mode) for women and for individuals whose trips originate and/or end in a large city. The most policy relevant difference is the lower cost sensitivity of travelers suggested by the COVNL model compared to the other two models. The implied cost of in-vehicle travel time is $17.20 per hour in the multinomial logit, $18.07 per hour in the standard nested logit model, and $23.70 per hour in the COVL model. The corresponding figures for out-of-vehicle travel time (computed at the mean trip distance value of 350.42 km) are $64.03 for the multinomial logit, $68.50 for the standard nested logit, and $87.38 for the COVNL model.

The parameters representing covariance heterogeneity indicate that there are significant differences in the correlation between car and train utilities among individuals. A positive parameter on a variable indicates that the variable increases the variance of the random utility components for the car and train modes conditional on a ground mode being chosen and, therefore, reduces the correlation between car and train utilities [see eqns (1) and (2)]. Higher income earnings reduces the correlation between car and train utilities.

*As one would expect, in the NL models with car and train grouped together and air and train grouped together, the implied estimate of the logsum parameter from the estimated α value in the transformed specification was the same as the estimated value of the logsum parameter in the untransformed specification.

Table 2. Intercity mode choice estimation results

Parm. type	Variable	Multinominal logit		NL model with car and train grouped		COVNL model with car and train grouped	
		Parm.	t-stat.	Parm.	t-stat.	Parm.	t-stat.
Mode Choice	Mode constants						
	Train	-0.4096	-0.99	0.2390	0.62	0.5996	1.53
	Air	-0.6041	-1.03	-0.8788	-1.64	-1.2810	-2.25
	Income ($\times 10^3$)						
	Train	-0.0007	-0.17	-0.0008	-0.27	-0.0022	-0.55
	Air	0.0369	8.95	0.0363	9.06	0.0362	9.07
	Traveling alone						
	Train	0.2672	1.69	0.2078	1.66	0.0539	0.35
	Air	0.3274	2.36	0.2985	2.21	0.3070	2.29
	Sex of traveler						
	Train	1.2200	7.57	0.9367	5.97	0.8741	5.86
	Air	0.5984	3.52	0.4773	2.84	0.4445	2.67
	Large city indicator						
	Train	1.1546	6.03	0.8280	4.58	0.7798	4.58
	Air	0.7738	4.47	0.6750	3.97	0.6577	3.89
	Frequency of service	0.0817	15.25	0.0826	16.11	0.0825	16.06
	Travel cost	-0.0349	-7.78	-0.0322	-7.40	0.0256	-5.74
	Travel time						
	Total	-0.0100	-11.01	-0.0097	-10.96	-0.0101	-11.23
	OVTT/Logdist	-0.1596	-8.97	-0.1586	-10.13	-0.1592	-10.12
Covar. Heterog.	Constant*	—	—	1.1505	—	-5.5371	—
	Income ($\times 10^3$)	—	—	—	—	0.0046	1.62
	Traveling alone	—	—	—	—	0.6897	2.27
	Log (trip distance)	—	—	—	—	1.0624	2.46
Log-likelihood at convergence†		-1471.40		-1468.35		-1462.71	
Adjusted l'hood ratio index		0.4182		0.4190		0.4200	

*The constant is implicitly restricted to positive infinity in the multinominal logit model so that the logsum parameter equals 1. The constant in the standard nested logit model implies a logsum parameter of $[1/(1+\exp(-1.1505))] = 0.7596$. This is the same value obtained if we estimate the logsum parameter directly. The standard error of the logsum parameter is 0.102; the t-statistic for the difference of the logsum parameter from 1 is 2.36. We do not report any t-statistics for the constant in the NL and COVNL models because the only reasonable test of the constant parameter would be against a value of positive infinity.
†The log-likelihood value at zero is -2970.65 and the log-likelihood value with only alternative specific constants and an IID error covariance matrix is -2549.53.

We provided an interpretation for this result earlier based on the larger accessibility of air as an alternative mode for high income travelers. Similar explanations can be provided for the reduced sensitivity between the ground modes for individuals traveling alone and individuals traveling long distances. For example, the unobserved perception of comfort and convenience of the ground modes might vary substantially more among ground-mode users traveling long distances than those traveling short distances. This leads to a reduced correlation between car and train utilities for large trip distances. Equivalently, car and train are closer substitutes for short trip distances than for long trip distances.

A statistical comparison of the multinomial logit model with the standard nested logit model leads to the rejection of the multinomial logit. A further likelihood ratio test between the NL and COVNL models results in the clear rejection of the hypothesis of covariance homogeneity (the likelihood ratio test value is 11.28 which is larger than the chi-squared statistic with 3 d.f. at any reasonable level of significance). Thus, both the standard nested logit model and multinomial logit models are mis-specified. Table 2 also evaluates the models in terms of the adjusted likelihood ratio index ($\bar{\rho}^2$).* These values again indicate that the COVNL model offers the best fit to the data.

*The adjusted likelihood ratio index is defined as follows: $\bar{\rho}^2 = 1 - \dfrac{L(M) - K}{L(C)}$ where $L(M)$ is the model log-likelihood value, $L(C)$ is the log-likelihood value with only alternative specific constants and an IID error covariance matrix, and K is the number of parameters (besides the alternative specific constants) in the model.

3.3. Policy implications for predicting rail mode shares

The objective of the original study for which the data were collected in the Toronto–Montreal corridor was to examine the effect of alternative improvements in rail level of service characteristics (see KPMG Peat Marwick & Koppelman, 1990). Consequently, we focus on the examination of the effects of policies which are associated with changes in rail level of service. This is achieved by computing the aggregate-level self and cross-elasticities of the effect of changes in rail level-of-service. These aggregate elasticities provide the proportional change in the expected market shares of each mode in response to an incremental change in a rail level-of-service variable. Ben-Akiva and Lerman (1985, p. 113) present the expressions for computing aggregate-level elasticities from individual-level elasticities. Forinash and Koppelman (1993) provide the expressions for the individual-level elasticities in the standard nested logit (NL) model. The expressions for the individual-level elasticities in the COVNL model are the same as those for the NL model except that the logsum parameter for each individual is reparameterized as in eqn (4).

Tables 3–5 show the elasticity effects of incremental improvements in rail level-of-service. The rail self-elasticities in Table 3 indicate that reductions in travel time, particularly the out-of-vehicle component, are the most effective means of increasing rail market share. The rail self-elasticity effects from the NL and COVNL models for improvements in rail frequency and travel times are larger than those obtained from the multinomial logit. Between the two nested logit models, the COVNL model indicates higher frequency and time self-elasticities than the standard nested model. Thus, the multinomial logit to a large extent, and the standard nested model to a smaller extent, underestimate the increase in rail market share in response to improvements in rail frequency and travel time service. The nested logit model with covariance heterogeneity indicates a substantially smaller rail cost self-elasticity than the other two models; that is, both the multinomial logit and standard nested logit models overestimate the sensitivity

Table 3. Aggregate rail self-elasticity matrix in response to improvement in rail service

	Model		
Rail level-of-service attribute	MNL	NL	COVNL
Increase in frequency	0.302	0.384	0.396
Decrease in cost	1.473	1.575	1.152
Decrease in in-vehicle time	1.539	1.804	1.928
Decrease in out-of-vehicle time	2.297	2.770	2.864

Table 4. Aggregate car cross-elasticity matrix in response to improvement in rail service

	Model		
Rail level-of-service attribute	MNL	NL	COVNL
Increase in frequency	–0.042	–0.056	–0.058
Decrease in cost	–0.159	–0.189	–0.149
Decrease in in-vehicle time	–0.169	–0.222	–0.234
Decrease in out-of-vehicle time	–0.303	–0.391	–0.403

Table 5. Aggregate air cross-elasticity matrix in response to improvement in rail service

	Model		
Rail level-of-service attribute	MNL	NL	COVNL
Increase in frequency	–0.025	–0.026	–0.027
Decrease in cost	–0.153	–0.136	–0.116
Decrease in in-vehicle time	–0.185	–0.180	–0.194
Decrease in out-of-vehicle time	–0.208	–0.207	–0.214

of rail market share to changes in rail cost. In summary, using the rail self-elasticities from the MNL and NL models will make a policy analyst much more conservative than (s)he should be in pursuing non-cost rail service improvements with a concomitant increase in rail travel fares.

The car cross-elasticities in Table 4 show, as expected, that improvements in rail frequency and travel times draw a larger market share from the car mode in the case of the nested models *vis-à-vis* the multinomial logit model. The cost cross-elasticity is lower in the COVNL model because of the overall lower sensitivity of mode utilities to cost in this model. The air cross-elasticities in Table 5 do not show much variation among the three models, except for the cost cross-elasticity which is lower in the COVNL model. In summary, the cross-elasticities suggest that improvements in rail frequency and travel time will alleviate auto congestion on highways more so than that estimated by the multinomial logit model. However, a policy aimed at reducing rail fares will be much less successful in alleviating auto congestion than as predicted by the standard nested logit model and, to a lesser extent, by the multinomial logit. Reducing rail fares will also have a much smaller impact on air market share than as suggested by the standard nested and multinomial logit models.

4. SUMMARY AND CONCLUSIONS

We have formulated and estimated a nested logit model which accommodates differential levels of sensitivity among nested alternatives across individuals. The resulting nested logit model with covariance heterogeneity (COVNL model) generalizes the extant standard nested logit (NL) model which imposes an equal correlation restriction among utilities of nested alternatives across all individuals. We use a full-information maximum likelihood method for estimation of the COVNL model.

The empirical analysis of the paper applied the multinomial logit, the nested logit, and the proposed COVNL formulations to the estimation of weekday, business, intercity travel mode choice in the Toronto–Montreal corridor. The results of the standard nested and COVNL formulations indicate that the nested structure with car and train (ground modes) grouped together and competing against the air mode is the most appropriate nesting structure.

An examination of the mode choice parameters from the multinomial logit, the nested logit, and the COVNL model show that all the three models provide estimates which are reasonable (it will be implied that the nested logit and COVNL models correspond to the structure with car and train grouped together). However, there are differences among the estimates in the three models. The most important of these is the lower cost sensitivity estimated by the COVNL model relative to the other two models. As a result, the COVNL model suggests a higher value of time than the other models.

The parameters representing covariance heterogeneity in the COVNL model indicate that there are significant differences in the correlation between car and train utilities among individuals. A statistical likelihood ratio test between the standard nested logit and the COVNL model rejects the hypothesis of covariance homogeneity at any reasonable level of significance. The standard nested and the COVNL models also reject the multinomial logit model. Thus, the multinomial logit and the standard nested logit models are mis-specified. Our results from the COVNL model indicate a lower sensitivity (or cross-elasticity) between train and car modes for individuals with high incomes, who travel alone, and who travel long distances.

The substantive implications of the effects of rail service improvements on mode market shares, expressed in terms of aggregate self- and cross-elasticities of rail service changes, are quite different among the multinomial logit, the nested logit, and the COVNL models. The multinomial logit and the standard nested logit underestimate the increase in rail market share in response to improvements in rail frequency and travel time service. These two models overestimate the sensitivity of rail market share to changes in rail cost. The cross-elasticities also indicate differences among the three models.

The standard nested and multinomial logit models overestimate the decrease in car and air market shares in response to a decrease in rail cost.

Overall, the empirical results point to the potential shortcomings of using the multinomial logit model or the standard nested logit model, and to the need for accommodating covariance heterogeneity across individuals in a nested model, in evaluating the impacts of policy actions aimed at alleviating intercity travel congestion.

REFERENCES

Ashenfelter O. C. and Layard R. (1986) *Handbook of Labor Economics.* North Holland, Amsterdam.

Ben-Akiva M. and Lerman S. R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand.* The MIT Press, Cambridge, MA.

Bhat C. R. (1995) A heteroscedastic extreme-value model of travel mode choice. *Transpn Res.* **29B**, 471–483.

Brownstone D. and Small K. A. (1989) Efficient estimation of nested logit models. *J. Econ. Bus. Stats.* 7, 67—74.

Bucklin R. E., Gupta S. and Han S. (1995) A brand's eye view of response segmentation in consumer brand choice behavior. *J. Market. Res.* **32**, 66–74.

Börsch-Supan A. (1987) *Econometric Analysis of Discrete Choice*, Lecture Notes in Economics and Mathematical Systems, p. 34. Springer-Verlag, Berlin.

Börsch-Supan A. (1990) Recent developments in flexible discrete choice models: nested logit analysis versus simulated moments probit analysis. *Spatial Choices and Processes* (Fischer M. M., Nijkamp P. and Papageorgiou Y. Y. Eds). Elsevier Science,

Börsch-Supan A. and Hajivassiliou V. A. (1993) Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *J. Econom.* **58**, 347–368.

Chamberlain G. (1980) Analysis of covariance with qualitative data. *Rev. Econ. Stud.* **47**, 225–238.

Corres S., Hajivassiliou V. A. and Ioannides Y. M. (1993) An empirical analysis of the dynamics of qualitative decisions of firms, Technical Report. Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Currim I. S. (1982) Predictive testing of consumer choice models not subject to independence of irrelevant alternative. *J. Market. Res.* **19**, 208–222.

Daly A. and Zachary S. (1979) Improved multiple choice models. *Identifying and Measuring the Determinants of Mode Choice* (Hensher D. and Dalvi Q. Eds). Teakfield, London.

Evers G. H. M. (1990) The residential location and workplace choice: a nested multinomial logit model. *Spatial Choices and Processes* (Fischer M. M., Nijkamp P. and Papageorgiou Y. Y. Eds). pp. 313–329. Elsevier Science/North Holland, Amsterdam.

Fischer M. M. and Aufhauser E. (1988) Housing choice in a regulated market: a nested multinomial logit analysis. *Geog. Anal.* **20**, 47–69.

Forinash C. V. and Koppelman F. S. (1993) Application and interpretation of nested logit models of intercity mode choice. *Transpn Res. Rec.* **1413**, 98–106.

Hausman J. and McFadden D. (1984) Specification tests for the multinomial logit model. *Econometrica* **52**, 1219–1240.

Hensher D. A., Smith N. C., Milthorpe F. M. and Barnard P. O. (1992) *Dimensions of Automobile Demand: A Longitudinal Study of Automobile Ownership and Use.* North Holland, Amsterdam.

Horowitz J. L. (1991) Reconsidering the multinomial probit model. *Transpn Res.* **25B**, 433–438.

Jain D. C., Vilcassim N. J. and Chintagunta P. K. (1994). A random-coefficients logit brand-choice model applied to panel data. *J. Bus. Econ. Stat.* **12**, 317–328.

Johnson N. and Kotz S. (1970) *Distributions in Statistics: Continuous Univariate Distributions*, Chap. 21. John Wiley, New York.

KPMG Peat Marwick and Koppelman F. S. (1990) Proposal for analysis of the market demand for high speed rail in the Quebec/Ontario corridor. Submitted to Ontario/Quebec Rapid Train Task Force.

KPMG Peat Marwick in association with ICF Kaiser Engineers, Inc., Midwest System Sciences, Resource Systems Group, Comsis Corporation and Transportation Consulting Group (1993) Florida High Speed and Intercity Rail Market and Ridership Study: Final Report, submitted to Florida Department of Transportation, July.

Manski C. and Lerman S. (1977) The estimation of choice probabilities from choice-based samples. *Econometrica* **45**.

McFadden D. (1973) Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics* (Zoremmbka P. Ed). Academic Press, New York.

McFadden D. (1978) Modelling the choice of residential location. *Spatial Interaction Theory and Residential Location* (Korlquist A. Ed), pp. 75–96. North Holland, Amsterdam.

Moon F. B. (1991) Houston moves forward again. *Civil Engng.* **61**, 40–43.

Nijkamp P. (1986) *Handbook of Regional and Urban Economics.* North Holland, Amsterdam.

Revelt D. and Train K. (1995) Incentives for appliance efficiency in a competitive energy industry, Technical Report. Department of Economics, University of California, Berkeley, CA.

Stopher P. R., Meyburg A. H. and Brög W. (1981) Travel behavior research: a perspective. *New Horizons in Travel Behavior Research* (Stopher P. R., Meyburg A. H. and Brög W. Eds). Lexington Books, Lexington, MA.

Waddell P. (1993) Exogenous workplace choice in residential location models; is the assumption valid? *Geog. Anal.* **25**, 65–82.