Paper No. 02-3882

Adaptation of Logit Kernel to Route Choice Situation

Shlomo Bekhor, Moshe E. Ben-Akiva, and M. Scott Ramming

An adaptation of the general logit kernel (LK) model to the route choice context is presented. Recent intelligent transportation systems applications have highlighted the need for better models of the behavioral processes involved in route choice. Several route choice models have been developed recently. However, few studies concentrated on the model estimation and applications for large urban networks. How LK can be adapted to route choice situations by suitably defining the elements of the model is described. The model is estimated using a sample formed from a route choice survey combined with network variables. Preliminary estimation results are presented and discussed.

Recent intelligent transportation systems applications have highlighted the need for better models of the behavioral processes involved in route choice. In particular, the desire to provide route guidance based on real-time traffic information to drivers highlights the fact that drivers have imperfect knowledge of traffic conditions and limited information-processing abilities. Given these limitations, it is not surprising to observe drivers making suboptimal (from the individual point of view) route choices. Drivers also exhibit a wide range of knowledge of network topology and route selection criteria, such as minimizing time or stress or maximizing the aesthetic experience of a trip.

Several route choice models, which are briefly described in the next section, were recently developed. Few studies, however, concentrated on the model estimation and applications for large urban networks.

This paper has two main purposes: first, to present an adaptation of the general logit kernel (LK) model to a route choice situation, and second, to present the estimation of the model for a real-size problem.

The paper is organized as follows: First, a brief review of route choice models is presented. Next, the paper presents the general LK model formulation and adapts it to route choice by suitably defining the elements of the model. A route choice survey conducted in the Boston area is used to estimate the model coefficients.

ROUTE CHOICE MODELS

The deterministic shortest path is the simplest route choice model. It is used in deterministic traffic assignment models, for example, all-or-nothing loading, and may be used as a component of user equilibrium assignment. The multinomial logit (MNL) and probit models were proposed long ago as generalizations to the deterministic model. Probit is based on the normal (or Gaussian) distribution and, thus, requires simulation. In comparison, MNL is based on the Gumbel distribution and has a well-known analytical form. The MNL model, however, is not suitable to model route choice, because it cannot account for similarities among routes.

Several types of models have recently been proposed to overcome the MNL drawbacks. These models represent modifications or generalizations of the logit structure. C-logit, proposed by Cascetta et al. (*1*) and path-size logit (PSL) presented in Ben-Akiva and Ramming (*2*) may be considered modifications to the MNL model because they add a correction term to path utilities but maintain the MNL model structure. Thus, they can be estimated using existing logit software.

The cross-nested logit model of Vovsha (*3*) and the paired combinatorial logit (PCL) model of Chu (*4*) were adapted for route choice in Prashker and Bekhor (*5*). Gliebe et al. (*6*) also adapted the PCL model for route choice. These models have a more general (and therefore more complex) error structure. They are members of the generalized extreme value (GEV) family of models developed by McFadden (*7*), which also includes MNL and nested logit.

PSL Model

The PSL model is similar to C-logit in that a correction term is added to a path's utility. However, PSL has a different theoretical basis. The notion of size comes from the theory of aggregate alternatives, which was first employed for destination and residence choice. However, unlike destination choice, in which zones may have a size representing thousands of elemental destinations (e.g., jobs), the largest size a path may have is one. Such a path shares no links with other paths (note that special treatment may be required for centroid connectors) and may be called a distinct or disjoint path.

The log of the path size is added to the path utility to form the PSL model from MNL:

$$P(i \mid C_n) = \frac{e^{V_{in} + \ln PS_{in}}}{\sum_{j \in C_n} e^{V_{jn} + \ln PS_{jn}}} = \frac{PS_{in} e^{V_{in}}}{\sum_{j \in C_n} PS_{jn} e^{V_{jn}}} \quad (1)$$

where

V_{in} = systematic utility of path i for person n,
C_n = path set for person n, and
PS_{in} = size of path i for person n.

Note that for a distinct path, the log of its size of one is zero, resulting in no utility adjustment. The extreme case of two paths

S. Bekhor, Department of Civil Engineering, Technion–Israel Institute of Technology, Technion Campus, Haifa, 32000, Israel. M. E. Ben-Akiva, Intelligent Transportation Systems Program, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307. M. S. Ramming, J. F. Sato and Associates, 5898 South Rapp Street, Littleton, CO 80120.

being created by "duplicating" or "splitting an existing path down the middle" results in each having a size of one-half. The path-size term may be calculated on the basis of the length of links in a path and the relative lengths of the paths that share a link. This calculation, therefore, is dependent on the specification of the choice set.

The path size can be specified in different ways, as investigated by Ramming (*8*). In this paper, we use the following path-size formulation:

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \frac{L_i^{\gamma}}{L_j^{\gamma}} \delta_{aj}} \qquad (2)$$

where

l_a = length of link a,
L_i = length of path i,
δ_{aj} = 1 if link a is in path j and 0 otherwise,
Γ_i = set of links of path i, and
γ = parameter to be calibrated.

The MNL and PSL models will be used in this paper as bases for comparison with the LK models presented in the next section.

Probit and LK Models

Because the independence of irrelevant alternatives–property of logit makes it difficult to represent the effect of overlapping paths, some researchers have examined the suitability of the probit model for route choice. Because the probit model is based on error terms having a multivariate normal distribution—as opposed to a Type I extreme value or Gumbel distribution as assumed in MNL and other GEV models—an arbitrary covariance structure may be specified. Daganzo (*9*) was one of the first to use the multinomial probit model.

In Japan, Yai et al. (*10*) provide a recent example of an application of the probit route choice model. The authors assume the covariance of route utilities is proportional to overlap length. Routes are also assumed to have heteroscedastic error terms in which variance is proportional to route length or impedance.

The difficulty in implementing the probit model is that no closed form exists for the Gaussian cumulative distribution function, so numerical techniques must be used. Numerical integration techniques are computationally feasible when the number of Gaussian variables (generally, the number of alternatives less one, which is normalized to be the base alternative) is small. Hajivassiliou et al. (*11*) present some alternative estimation methods, and Bolduc (*12*) advises that maximum simulated likelihood estimation with a Geweke-Hajivassiliou-Keane probability simulator is the preferred method for transportation modeling with large samples and choice sets.

Choice models with combinations of Gaussian and Type I extreme value error terms have been proposed by researchers such as McFadden and Train (*13*), who call the resulting model mixed logit, and by Ben-Akiva and Bolduc (*14*), who refer to the resulting system as multinomial probit with LK, or simply LK. Other authors may refer to this formulation as hybrid logit.

These combined models suffer from the same computational difficulties as pure multinomial probit. Programs to estimate these types of models are widely available [see Train et al. (*15*)]. Bhat (*16*) presented the use of intelligent drawing mechanisms (in many cases nonrandom draws known as Halton sequences). These quasi-random draws are designed to cover the integration space in a more uniform way and, therefore, can significantly reduce the number of draws required.

Recently, Han et al. (*17*) used the mixed logit model to investigate taste heterogeneity across drivers and the possible correlation between repeated choices. The data set used was a stated preference (SP) survey conducted by the Swedish Transport Research Institute. The number of alternatives was therefore limited by the SP experiment.

Paag et al. (*18*) and Nielsen et al. (*19*) estimated route choice models for the Harbour Tunnel project in Copenhagen. The data source for this project was also collected from an SP survey. The models that included normally distributed coefficients led to better results than the models without distributed coefficients. The route choice models were applied in a multiclass stochastic assignment model.

LK MODEL FORMULATION

In the LK model, the random utility term is made up of two components: a probit-like component with a multivariate distribution and an independent and identically distributed (i.i.d.) Gumbel random variate.

The probit-like term captures the interdependencies among the alternatives. We specify these interdependencies using a factor analytic structure, which is a flexible specification that accommodates different error structures, as was shown by Walker (*20*). It also has the ability to capture complex covariance structures with relatively few parameters.

The general form of the factor analytic LK model (in vector notation) is presented following Ben-Akiva et al. (*21*). For simplicity, we omit the index n (person).

$$\mathbf{U} = \boldsymbol{\beta}^T\mathbf{X} + \epsilon = \boldsymbol{\beta}^T\mathbf{X} + \mathbf{F}\boldsymbol{\xi} + \mathrm{j} + \nu \qquad (3)$$

where

$\mathbf{U}$ = (J*1) vector of utilities,
$\boldsymbol{\beta}$ = (K*1) vector of unknown parameters,
$\mathbf{X}$ = (J*K) matrix of explanatory variables,
$\mathbf{F}$ = (J*M) factor loadings matrix, and
$\boldsymbol{\xi}$ = (M*1) vector of m multivariate distributed latent factors.

In the LK model, the random utility term ϵ is made up of two components: a probit-like component with a multivariate distribution $\mathbf{F}\boldsymbol{\xi}$ and an i.i.d. Gumbel random variate ν. The probit-like term captures the interdependencies among the alternatives. We specify these interdependencies using a factor analytic structure, which is a flexible specification. It has the ability of capturing complex covariance structures with relatively few parameters. $\mathbf{F}$ is a matrix of the factor loadings that map the factors to the error vector.

For estimation, it is desirable to specify the factors such that they are independent, and we therefore decompose $\boldsymbol{\xi}$ as follows:

$$\boldsymbol{j} = \mathbf{T}\boldsymbol{\zeta} \qquad (4)$$

where $\mathbf{T}$ is the (M*M) lower triangular matrix of unknown parameters, and $\boldsymbol{\zeta}$ is the (M*1) vector of i.i.d. random variables with zero mean and unit variance.

$\mathbf{TT}^T$ is the covariance matrix of $\boldsymbol{\xi}$, and $\mathbf{T}$ is the Cholesky factorization of it. The number of factors M can be less than, equal to, or greater than the number of alternatives. In this paper it is assumed that the factors $\boldsymbol{\zeta}$ have standard normal distributions. Given these assumptions, the covariance matrix is given by the following expression:

$$\text{cov}(\boldsymbol{\epsilon}) = \mathbf{FTT}^T\mathbf{F}^T + (g/\mu^2)\mathbf{I} \quad (5)$$

where

g = variance of a standard Gumbel variable ($\pi^2 / 6$),
ν = (J*1) vector of i.i.d. Gumbel variables with scale parameter μ, and
$\mathbf{I}$ = identity matrix.

The elements of **F** and **T** may be estimated or specified from data. As presented later in the paper, we will define the **F** matrix in a convenient way to adapt it for route choice.

To obtain the probability function, we make use of the convenient logit formulation as follows. If the factors $\boldsymbol{\zeta}$ are known, the following expression is obtained:

$$\Lambda(i|\boldsymbol{\zeta}) = \frac{\exp[\mu(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{F}_i\mathbf{T}\boldsymbol{\zeta})]}{\sum_j \exp[\mu(\mathbf{X}_j\boldsymbol{\beta} + \mathbf{F}_j\mathbf{T}\boldsymbol{\zeta})]} \quad (6)$$

where $\Lambda(i|\boldsymbol{\zeta})$ is the probability to choose alternative i given $\boldsymbol{\zeta}$.

The function above is equivalent to the MNL formulation. Because the factors are unknown, the unconditional probability is given by the following:

$$P(i) = \int_{\text{j}} \Lambda(i|\zeta)\prod_{m=1}^{M} \phi(\boldsymbol{\zeta})d\boldsymbol{\zeta} \quad (7)$$

where $\phi(\boldsymbol{\zeta})$ is the standard univariate normal density function.

The advantage of the LK is that we can estimate the probability function by simulation:

$$P(i) = \frac{1}{D}\sum_{d=1}^{D} \Lambda(i|\boldsymbol{\zeta}^d) \quad (8)$$

where D is the number of simulation draws, and $\boldsymbol{\zeta}^d$ denotes draw d from the distribution of $\boldsymbol{\zeta}$.

LK ROUTE CHOICE MODEL

The key issue in applying the LK model to route choice is the specification of the elements of the **F** and **T** matrices. One reasonable specification of the LK model for route choice would be to assume that the covariance of path utilities is proportional to the length by which paths overlap. This is a common assumption used in the probit model; see for example, Daganzo and Sheffi (*22*) or, more recently, Yai et al. (*10*).

The following assumptions are required to adapt LK to route choice:

- Link-specific factors are i.i.d. normal,
- Variance is proportional to the link length,
- The **F** matrix is the link-path incidence matrix, and
- The **T** matrix is the link factors variance matrix (a diagonal matrix).

To illustrate the adaptation of the LK to route choice, we consider the well-known “red bus–blue bus” network. Figure 1 shows the example network.

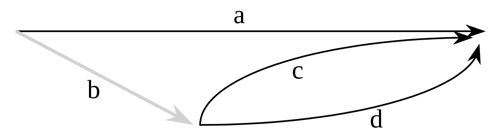

FIGURE 1 Red bus–blue bus network: Route 1 = link a, Route 2 = links b–c, Route 3 = links b–d.

This network is composed of three routes. Two routes share a common link (Link b in the figure). The correspondent **F** matrix is defined as the link-path incidence matrix as follows:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad (9)$$

The dimensions of the matrix are three rows, corresponding to the routes, and four columns, corresponding to the links.

The triangular matrix is obtained assuming the variance proportional to the link free-flow travel costs as follows:

$$\mathbf{T} = \begin{bmatrix} \sigma_a & 0 & 0 & 0 \\ 0 & \sigma_b & 0 & 0 \\ 0 & 0 & \sigma_c & 0 \\ 0 & 0 & 0 & \sigma_d \end{bmatrix} = \sigma \begin{bmatrix} \sqrt{t_a} & 0 & 0 & 0 \\ 0 & \sqrt{t_b} & 0 & 0 \\ 0 & 0 & \sqrt{t_c} & 0 \\ 0 & 0 & 0 & \sqrt{t_d} \end{bmatrix} \quad (10)$$

This matrix has only positive elements in the diagonal, which allows quick computation of the utility vector and covariance matrix. Note that there is only one parameter to estimate. The utility vector is then obtained as follows:

$$\mathbf{U} = \begin{bmatrix} \boldsymbol{\beta}\mathbf{X}_1 + \sigma_a\zeta_1 + \nu_1 \\ \boldsymbol{\beta}\mathbf{X}_2 + \sigma_b\zeta_2 + \sigma_c\zeta_3 + \nu_2 \\ \boldsymbol{\beta}\mathbf{X}_3 + \sigma_b\zeta_2 + \sigma_d\zeta_4 + \nu_3 \end{bmatrix} \quad (11)$$

The covariance matrix is obtained as follows:

$$\mathbf{FTT}^T\mathbf{F}^T = \begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_b^2 + \sigma_c^2 & \sigma_b^2 \\ 0 & \sigma_b^2 & \sigma_b^2 + \sigma_d^2 \end{bmatrix} \quad (12)$$

This concludes the adaptation of the LK formulation to route choice. The next section presents estimation results for a test case.

ESTIMATION RESULTS

Route choice data come from a *1997 Transportation Survey of Faculty and Staff* conducted by the MIT Planning Office. Drivers were asked to provide a written description of their habitual routes, as they would describe it to a neighbor or colleague. When route descriptions contained gaps, we used the shortest path to connect known portions of the survey respondent’s route. We omitted observations in which the respondent made stops along the way or did not provide enough information from which to construct a coherent route.

A total of 188 observations were considered after the cleaning process. Ramming (*8*) presents a detailed description of the sample method and characteristics. It is important to stress that all observations refer to MIT staff (not students); 90% of the respondents were full-time employees.

Choice Set Generation

To estimate the choice models, we first need to generate a choice set. We have performed route generation experiments using a highway network database developed by central transportation planning staff (CTPS), the Metropolitan Planning Organization for the Boston region. The highway network covers an area of approximately 2,800 mi^2 where about 4 million inhabitants reside. The network consists of more than 800 zones, about 13,000 nodes, and about 34,000 one-way links.

There are several link attributes in the database: distance, free-flow time, estimated or congested time (i.e., the output of the CTPS traffic assignment model), capacity, number of lanes, tolls, assigned volume, and functional class. In addition, the database includes the presence of government-numbered signage (e.g., Interstate 93, US-1, or MA-16) and indicators of security such as neighborhood income and employment.

Generation of the choice set is described in detail in Bekhor et al. (*23*). The method used to generate the routes was to draw a sample of link impedances on the basis of the congested link travel times and find the shortest path for each observation. This procedure was repeated for 48 draws. Deterministic shortest paths that minimize distance, free-flow time, and congested time were also added, allowing each observation to have a choice set containing up to 51 alternatives.

For 159 (out of the original 188) observations, the overlap between the observed route and the closest-matching generated route exceeded the prespecified threshold of 80%. Bekhor et al. (*23*) argued that for the Boston data set, this threshold is acceptable, given the imperfections in the network data and survey method.

Distribution of the number of alternative routes in the choice set of the 159 observations considered for model estimation is presented in Figure 2. A few observations consider only two alternative routes. About 80% of the observations have a choice set of at least 20 routes; the median choice set size is about 30 routes. For a few observations, the 51 draws in the final route set generation procedure yielded 51 unique paths. Recall that in the LK route choice formulation, the number of links in the choice set determines the number of unknown factors to be simulated and, thus, affects the computational requirements for estimation.

The variety of respondents' choice sets is further illustrated in Figure 3, which considers the size of choice sets in terms of links, rather than paths. Thus, a more distant origin will have a larger number of links in its choice set for two reasons: (*a*) because it is more distant, drivers will need to traverse more links to reach their destination; and (*b*) because of the greater distance between origin and destination, drivers will likely have more alternative paths available. The smallest choice set contains about 20 links, and the largest contains about 850. The median choice set size is about 250 links.

MNL and PSL Estimation Results

Table 1 presents the best estimates obtained for the MNL and PSL route choice models. These models are estimated to provide a basis for comparison with the LK model. Parameter estimates are shown in bold, followed by *t*-statistics. The *t*-statistics are for the hypothesis of a zero parameter value, except for the ln path-size term, which is calculated against a null hypothesis that the coefficient equals one.

We first discuss the level-of-service variables. The coefficients on distance and free-flow time are all significant and negative, as expected. The best estimates of travel time delay were obtained for a nonlinear (logarithmic) form. Note that these estimates are segmented by income.

The coefficient on the Tobin Bridge dummy is positive and significant. The Massachusetts Turnpike and Sumner Tunnel dummy coefficients are not always significantly different from zero. We prefer to retain the coefficients on three toll facility dummies even when they are not significantly different from zero. These facility

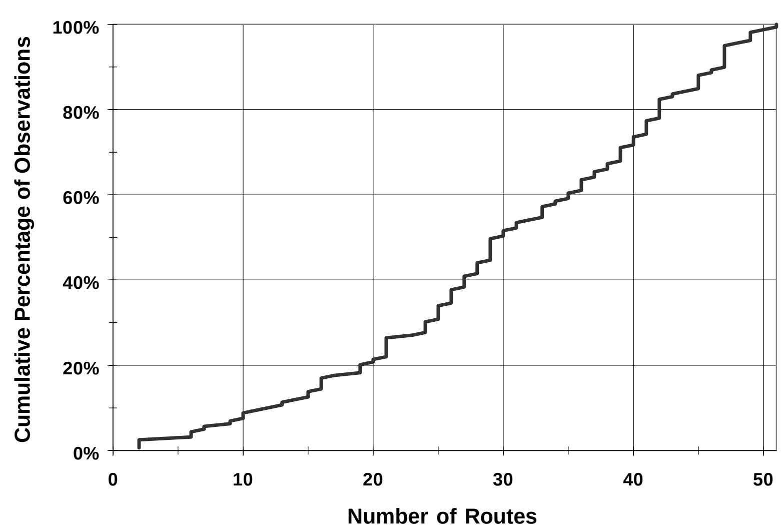

FIGURE 2 Cumulative distribution of choice set size.

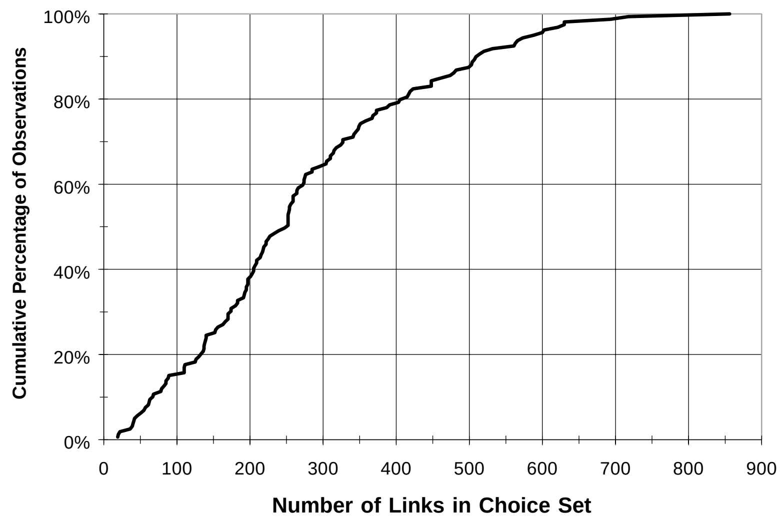

FIGURE 3 Distribution of links in choice sets.

dummies serve a role similar to that of alternative-specific constants in mode choice. Recall that the order of applying path generation algorithms is arbitrary.

The Massachusetts Turnpike dummy coefficient is negative and equivalent to about 1 min of free-flow time. The Tobin Bridge and Sumner Tunnel dummies are positive and equivalent to about 4 or 5 min of free-flow time. We interpret this as the Tobin Bridge and Sumner Tunnel being very prominent facilities for crossing a major barrier—the Boston Harbor and Mystic River. Travelers coming from northeast of MIT might use these prominent features in their trip planning. The Massachusetts Turnpike, being a linear feature rather than a barrier crossing, may not have the same prominence. Its negative coefficient may be more related to delays experienced at toll booths.

The coefficient on numbered route time is positive and significant. This suggests that drivers prefer to drive on signed highways, which provide more information about destinations and connecting roads. The label dummies indicating the least-distance and least-

TABLE 1 MNL and PSL Estimation Results

Variable	**MNL**	**PSL $\gamma = 1$**	**PSL $\gamma = 4$**	**PSL $\gamma = 99$**	**PSL $\gamma = \infty$**
1. Distance (miles)	**-0.253** -2.4	**-0.256** -2.4	**-0.248** -2.3	**-0.204** -2.0	**-0.212** -2.1
2. Free-Flow Time (minutes)	**-0.601** -6.6	**-0.669** -6.4	**-0.653** -6.7	**-0.492** -6.1	**-0.513** -6.3
3. Path Uses Mass. Pike (Dummy)	**-0.640** -0.9	**-0.830** -1.2	**-0.760** -1.1	**-0.520** -0.9	**-0.490** -0.8
4. Path Uses Tobin Bridge (Dummy)	**2.90** 3.1	**3.10** 3.3	**3.04** 3.2	**2.76** 3.1	**2.75** 3.1
5. Path Uses Sumner Tunnel (Dummy)	**2.18** 1.8	**1.89** 1.6	**2.09** 1.7	**1.73** 1.5	**1.92** 1.7
6. Ln(delay) for No Income Reported	**-5.13** -2.6	**-5.26** -2.7	**-5.25** -2.7	**-4.26** -2.5	**-4.45** -2.5
7. Ln(delay) for Income < $100,000 per year	**-0.205** -0.5	**-0.348** -0.8	**-0.325** -0.7	**-0.614** -1.5	**-0.583** -1.4
8. Ln(delay) for Income ≥ $100,000 per year	**-2.562** -2.7	**-2.719** -2.9	**-2.658** -2.8	**-2.736** -3.1	**-2.676** -3.0
9. Time spent on Government-Numbered Route	**0.112** 3.5	**0.110** 3.4	**0.113** 3.5	**0.089** 2.9	**0.090** 2.9
10. Path with Least Distance Label (Dummy)	**1.056** 4.2	**0.982** 3.8	**0.993** 3.9	**0.725** 2.9	**0.759** 3.0
11. Path with Least Estimated Time (Dummy)	**0.971** 4.3	**0.958** 4.2	**0.956** 4.2	**0.466** 1.9	**0.377** 1.5
12. Ln Path Size based on FF Time, $\gamma = \infty$		**0.567** -1.0	**0.498** -1.5	**0.933** -0.5	**0.730** -2.2
Number of Observations	159	159	159	159	159
Initial Log-Likelihood	-519.7	-519.7	-519.7	-519.7	-519.7
Final Log-Likelihood	-410.8	-409.9	-409.7	-397.7	-393.1
Number of Parameters	11	12	12	12	12
Rho-Bar Squared	0.188	0.188	0.189	0.212	0.221

estimated-time route are also positive and significant. These paths are likely to be well known among regular commuters.

The final point is related to the calibration of the path-size parameter γ. Note that the log-likelihood and rho-bar squared are monotonically increasing in γ. As discussed in Ramming (*8*), low values of γ may lead to counterintuitive results, so we did not expect good fit from the models estimated on these data sets. However, notice that there is a jump of about four log-likelihood points when we go from a gamma of 99 to infinity. This result is surprising, because it suggests the path-size term is based on a deterministic label. That is, a path that overlaps with an infinitesimally shorter one gets absolutely no link size for the overlapping segment (and therefore a substantially reduced path size). An empirical consideration is that the path-size label (here, free-flow time) must be known with certainty, or specifications with path size based on different labels must be tested. This result invites skepticism and suggests that researchers use a large finite value of gamma should calibrating the path size formulation be prohibitive.

LK Estimation Results

To estimate the LK model, we first need to prepare the factor analytic matrix, as illustrated by Equations 9 and 10. To save memory, only the links that were used in the choice set are stored. As many as 856 links per observation were selected in the choice set generation process, and therefore the maximum dimension of the **F** matrix is 51 (maximum number of routes for one observation) × 856 links.

We used the maximum simulated likelihood estimation method because of the large number of Gaussian variates (one per link in each respondent's choice set). It is not computationally feasible to estimate these model structures using numerical integration. Therefore, the number of draws used in the simulation and estimation process is shown at the top of each column. All the draws were performed using pseudorandom numbers. Walker (*20*) advises that the number of draws must be sufficiently large so that parameter results are "stable" or "robust" as the number of draws increases.

The coefficient estimates presented in the table have the same signs as those presented for the PSL models, so the general interpretation is unchanged. However, notice that the LK estimates do not appear to be the same multiple of the PSL coefficients, and that the significance of some coefficients has changed. For example, the coefficient on distance did not increase as much compared with the coefficient on free-flow time. Also, the distance coefficient was statistically different from zero in the PSL specification, whereas this assertion cannot be made for the LK specification. In contrast, the coefficients on the log of delay for households with annual incomes under $100,000 per year have higher *t*-statistics under the LK specification.

In the LK specifications with path size, the path-size coefficient is statistically different from zero, but not significantly different from one. The estimate of the Gaussian covariance parameter σ is also statistically different from zero. This suggests that the Gaussian covariances, which are proportional to path overlap free-flow times, capture a different effect than path size. Also notice that the LK specifications with path size have a better fit than the MNL and PSL models presented in Table 1.

The coefficient estimates for the LK model presented in Table 2 have a different scale or magnitude than the MNL and PSL coefficients shown in Table 1. Recall that for the LK model, the error term in the route utilities consists of a Gaussian term (which may be

TABLE 2 LK Estimation Results

	Number of Simulation Draws				
Variable	**10,000**	**4,096**	**10,000**	**24,649**	**100,000**
1. Distance (miles)	**-0.179** -0.4	**-0.274** -0.6	**-0.323** -0.7	**-0.327** -0.7	**-0.296** -0.7
2. Free-Flow Time (minutes)	**-4.00** -7.2	**-3.66** -4.2	**-3.44** -2.3	**-3.34** -1.7	**-3.43** -2.0
3. Path Uses Mass. Pike (Dummy)	**-5.96** -1.7	**-5.65** -1.6	**-5.40** -1.4	**-5.19** -1.2	**-5.41** -1.3
4. Path Uses Tobin Bridge (Dummy)	**6.72** 1.5	**8.58** 1.6	**7.82** 1.3	**8.02** 1.2	**8.39** 1.4
5. Path Uses Sumner Tunnel (Dummy)	**4.46** 0.6	**5.25** 0.7	**4.96** 0.6	**4.57** 0.6	**4.81** 0.7
6. Ln(delay) for No Income Reported	**-13.20** -2.0	**-13.50** -1.9	**-12.40** -1.6	**-12.50** -1.3	**-14.2** -1.5
7. Ln(delay) for Income <$100,000 per year	**-2.57** -1.5	**-3.06** -1.8	**-2.78** -1.5	**-2.82** -1.3	**-2.86** -1.4
8. Ln(delay) for Income ≥$100,000 per year	**-10.00** -1.8	**-10.80** -1.9	**-10.10** -1.6	**-10.00** -1.3	**-10.20** -1.5
9. Time spent on Government-Numbered Route	**0.378** 3.0	**0.367** 2.1	**0.338** 2.1	**0.335** 1.7	**0.345** 1.9
10. Path with Least Distance Label (Dummy)	**2.52** 3.2	**1.82** 2.1	**1.73** 1.8	**1.67** 1.5	**1.73** 1.7
11. Path with Least Estimated Time (Dummy)	**2.10** 2.9	**1.32** 1.8	**1.26** 1.6	**1.26** 1.5	**1.19** 1.5
12. Ln Path Size based on FF Time, $\gamma = \infty$		**1.13** 2.2	**0.99** 2.0	**1.07** 2.0	**1.03** 2.0
13. Gaussian covariance parameter based on Free-flow Time	**2.17** 5.8	**2.12** 3.7	**1.99** 2.3	**1.97** 1.7	**1.97** 2.0
Number of Observations	159	159	159	159	159
Initial Log-Likelihood	-519.7	-519.7	-519.7	-519.7	-519.7
Final Log-Likelihood	-390.0	-382.4	-382.4	-382.3	-382.0
Number of Parameters	12	13	13	13	13
Rho-Bar Squared	0.226	0.239	0.239	0.239	0.240

correlated with the corresponding terms for other routes) and an independent Gumbel term. In MNL and PSL models, route utilities have only the independent Gumbel term. Therefore, the Gumbel term of these models has a greater variance than the Gumbel term of the LK models. Because of the customary normalization of logit models, specifications with greater variance will yield utility parameters (β's) with lower magnitudes. Therefore, we expect the LK parameters to have a larger scale than the MNL coefficients.

Table 3 presents the LK estimates of Table 2 scaled so that the coefficient on free-flow time has the same magnitude as the PSL model of Table 1.

The scaled coefficients on distance, log of delay for high-income travelers, and label dummies have a much smaller magnitude than the corresponding PSL coefficients. However, notice that the path-size coefficient has a greater magnitude and is quite close to one.

DISCUSSION OF RESULTS

This paper showed how LK was adapted to model route choice for a large problem. The key issue is the assumption of a diagonal factor analytic matrix, which allowed computation of the covariance matrix at affordable computer resources.

Preliminary estimation results showed that the LK model had a better overall fit than simpler models such as MNL and PSL. However, these results should be verified for other data sets because the data used in this paper were based on a relatively small number of observations (159 in total). Results should also be compared against other closed-form models, such as C-logit, cross-nested logit, and PCL models.

In this paper we assumed that the choice set in all models was fixed and independent from the choice model. Because there is information only about the chosen route, the idea was to keep generating routes until all observations were covered (we decided to stop when for 159 observations we reached 80% overlap). This means that for a different data set (e.g., a sparse network such as interurban roads), the number of alternatives generated using the same method may be smaller. Estimation of the models for different choice set sizes is left for further research. Ideally, future route choice surveys would collect mental map data or other indicators of the awareness of alternative routes.

Another area of research is related to the stability of the results. As pointed out by Walker (*20*), there is a need to perform a large number of draws to obtain stable results. We suggest that for the dimensions of the problem presented in this paper, a very large number of draws is needed; after 100,000 draws, results were still not satisfactorily stable on the basis of the coefficient and standard error estimates. Recent results presented by Bhat (*16*) indicate that quasi-random numbers significantly reduce the number of draws required to obtain stability. This is especially important when the number of alternatives is large (as in the case of this paper), because of the large dimension of the integral.

Recently, Bhat (*24*) addressed two aspects of standard Halton draws that can limit their applicability. First, coverage of the integration domain deteriorates quite rapidly in higher dimensions. Second, standard Halton draws are deterministic, which prevents computation of statistical error bounds associated with simulation noise. To address these limitations, Bhat proposed the use of scrambled Halton draws to ensure good coverage in higher dimensions and a randomization technique to allow measurement of simulation variance.

We did not account for taste variation in this paper. Recent results showed that this is an important factor in route choice. In this paper we focused on the factor analytic formulation of the LK model and the adaptation for route choice. The inclusion of random coefficient specifications for the route choice model and comparison with results presented by other authors are left for further research.

TABLE 3 Comparison Between PSL and LK Estimation Results

Variable	PSL (Table 1)	LK (Table 2)	Scaled LK
1. Distance (miles)	**-0.212** -2.1	**-0.296** -0.7	**-0.044** -0.7
2. Free-Flow Time (minutes)	**-0.513** -6.3	**-3.43** -2.0	**-0.513** -2.0
3. Path Uses Mass. Pike (Dummy)	**-0.49** -0.8	**-5.41** -1.3	**-0.81** -1.3
4. Path Uses Tobin Bridge (Dummy)	**2.75** 3.1	**8.39** 1.4	**1.26** 1.4
5. Path Uses Sumner Tunnel (Dummy)	**1.92** 1.7	**4.81** 0.7	**0.72** 0.7
6. Ln(delay) for No Income Reported	**-4.45** -2.5	**-14.2** -1.5	**-2.12** -1.5
7. Ln(delay) for Income <\$100,000 per year	**-0.583** -1.4	**-2.86** -1.4	**-0.428** -1.4
8. Ln(delay) for Income ≥\$100,000 per year	**-2.676** -3.0	**-10.20** -1.5	**-1.526** -1.5
9. Time spent on Government-Numbered Route	**0.090** 2.9	**0.345** 1.9	**0.0516** 1.9
10. Path with Least Distance Label (Dummy)	**0.759** 3.0	**1.73** 1.7	**0.259** 1.7
11. Path with Least Estimated Time (Dummy)	**0.377** 1.5	**1.19** 1.5	**0.178** 1.5
12. Ln Path Size based on FF Time, $\gamma = \infty$	**0.730** -2.2	**1.03** 2.0	**1.03** 2.0
13. Gaussian covariance parameter based on Free-flow Time		**1.97** 2.0	**0.295** 2.0

REFERENCES

1. Cascetta, E., A. Nuzzolo, F. Russo, and A. Vitetta. A Modified Logit Route Choice Model Overcoming Path Overlapping Problems: Specification and Some Calibration Results for Interurban Networks. *Transportation and Traffic Theory: Proc., 13th International Symposium on Transportation and Traffic Theory* (J.-B. Lesort, ed.). Lyons, France. Pergamon, 1996.
2. Ben-Akiva, M. E., and M. S. Ramming. Lecture Notes: Discrete Choice Models of Traveler Behavior in Networks. Prepared for Advanced Methods for Planning and Management of Transportation Networks, Capri, Italy, 1998.
3. Vovsha, P. Application of Cross-Nested Logit Model to Mode Choice in the Tel-Aviv, Israel, Metropolitan Area. In *Transportation Research Record 1607*, TRB, National Research Council, Washington, D.C., 1997, pp. 6–15.
4. Chu, C. A Paired Combinatorial Logit Model for Travel Demand Analysis. *Proc., 5th World Conference on Transportation Research,* Vol. 4, Ventura, Calif., 1989, pp. 295–309.
5. Prashker, J. N., and S. Bekhor. Investigation of Stochastic Network Loading Procedures. In *Transportation Research Record 1645*, TRB, National Research Council, Washington, D.C., 1998, pp. 94–102.
6. Gliebe, J. P., F. S. Koppelman, and A. Ziliaskopoulos. Route Choice Using Paired Combinatorial Logit Model. Presented at 78th Annual Meeting of the Transportation Research Board, Washington, D.C., 1999.
7. McFadden, D. Modeling the Choice of Residential Location. In *Spatial Interaction Theory and Residential Location* (A. Karlqvist et al., eds.), North Holland, Amsterdam, Netherlands, 1978, 75–96.
8. Ramming, M. S. Network Knowledge and Route Choice. Unpublished Ph.D. thesis. Massachusetts Institute of Technology, Cambridge, 2001.
9. Daganzo, C. F. Multinomial Probit and Qualitative Choice: A Computationally Efficient Algorithm. *Transportation Science*, Vol. 11, 1977, pp. 338–358.

10. Yai, T., S. Iwakura, and S. Morichi. Multinomial Probit with Structured Covariance for Route Choice Behavior. *Transportation Research,* Vol. 31B, 1997, pp. 195–207.
11. Hajivassiliou, V. A., D. McFadden, and P. A. Ruud. Simulation of Multivariate Normal Rectangle Probabilities: Methods and Programs. *Journal of Economics,* Vol. 72, 1996, pp. 2–85.
12. Bolduc, D. A Practical Technique to Estimate Multinomial Probit Models in Transportation. *Transportation Research Part B,* Vol. 33B, No. 1, 1999, pp. 63–79.
13. McFadden, D., and K. Train. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics,* Vol. 15, No. 5, 2000 pp. 447–470.
14. Ben-Akiva, M., and D. Bolduc. Multinomial Probit with a Logit Kernel and a General Parametric Specification of the Covariance Structure. Working Paper. Massachusetts Institute of Technology, Cambridge, 1996.
15. Train, K., D. Revelt, and P. Ruud. Mixed Logit Estimation Routine for Cross-Sectional Data. elsa.berkeley.edu/Software/abstracts/train0196.html. 1999.
16. Bhat, C. R. Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research Part B,* Vol. 35, 2001, pp. 677–693.
17. Han, B., S. Algers, and L. Engelson. Accommodating Drivers' Taste Variation and Repeated Choice Correlation in Route Choice Modeling by Using Mixed Logit Model. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C., 2001.
18. Paag, H., A. Daly, and C. Rohr. Predicting Use of the Copenhagen Harbour Tunnel. *9th International Association for Travel Behaviour Research Conference* (CD-ROM), Gold Coast, Australia, 2000.
19. Nielsen, O. A., A. Daly, and R. D. Frederiksen. A Stochastic Multi-class Route Choice Model with Distributed Coefficients for Time and Cost. *9th International Association for Travel Behaviour Research Conference* (CD-ROM), Gold Coast, Australia, 2000.
20. Walker, J. *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables.* Ph.D. thesis. Massachusetts Institute of Technology, 2001.
21. Ben-Akiva, M., D. Bolduc, and J. Walker. Specification, Identification, and Estimation of the Logit Kernel (or Continuous Mixed Logit) Model. Working paper. Massachusetts Institute of Technology, Cambridge, 2001.
22. Daganzo, C. F., and Y. Sheffi. On Stochastic Models of Traffic Assignment. *Transportation Science*, Vol. 11, 1977, pp. 253–274.
23. Bekhor, S., M. Ben-Akiva, and M. S. Ramming. Estimating Route Choice Models for Large Urban Networks. *9th World Conference on Transport Research,* Seoul, Korea, 2001.
24. Bhat, C. R. Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. Working paper. University of Texas at Austin, 2001.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.