Logit kernel (or mixed logit) models for large multidimensional choice problems: identification and estimation

John L. Bowman, Ph. D., Research Affiliate, Massachusetts Institute of Technology, 5 Beals Street Apt. 3, Brookline, MA 02446, USA, voice: +1-617-232-3478, email: John_L_Bowman@alum.mit.edu, website: http://JBowman.net

originally submitted July 29, 2003
first revision submitted September 15, 2003

word count: 7221 + 9 tables = 9471

ABSTRACT

This paper presents an identification rule and insights for estimation of the class of error component logit kernel (EC) models—logit kernel (or mixed logit) models involving heteroscedasticity and subsets of alternatives with shared unobserved attributes. EC includes analogs of nested logit (NL), cross-nested logit (CNL), paired combinatorial logit (PCL), heteroscedastic logit (HL), their combinations, and other generalized extreme value (GEV) forms. The identification rule is necessary but not sufficient; however, it is simpler to use than the underlying necessary rank condition, and is adequate for complex model specifications. A case study demonstrates the specification, identification and estimation (using maximum simulated likelihood (MSL) with shuffled Halton draws) of the type of model for which EC is useful—one with large choice set and a choice outcome consisting of two or more variables considered simultaneously. It models a worker's day activity pattern—the choice among 162 alternatives for completing the non-work portion of the day—with each alternative defining a configuration of optional tours and commute stops for maintenance and discretionary purposes. The estimated EC is superior to a nested logit model, with significant covariance parameters for several overlapping subsets of the choice set. The case study also demonstrates (a) the importance of testing for simulation error and using many draws in MSL simulation (in this case at least 1500 draws for each simulated probability), and (b) the pending practicality of estimating EC models for large multidimensional choice problems.

Logit kernel (or mixed logit) models for large multidimensional choice problems: identification and estimation

John L. Bowman, Ph. D., Research Affiliate, Massachusetts Institute of Technology, 5 Beals Street Apt. 3, Brookline, MA 02446, USA, voice: +1-617-232-3478, email: John_L_Bowman@alum.mit.edu, website: http://JBowman.net

word count: 7221 + 9 tables = 9471

ABSTRACT

This paper presents an identification rule and insights for estimation of the class of error component logit kernel (EC) models—logit kernel (or mixed logit) models involving heteroscedasticity and subsets of alternatives with shared unobserved attributes. EC includes analogs of nested logit (NL), cross-nested logit (CNL), paired combinatorial logit (PCL), heteroscedastic logit (HL), their combinations, and other generalized extreme value (GEV) forms. The identification rule is necessary but not sufficient; however, it is simpler to use than the underlying necessary rank condition, and is adequate for complex model specifications. A case study demonstrates the specification, identification and estimation (using maximum simulated likelihood (MSL) with shuffled Halton draws) of the type of model for which EC is useful—one with large choice set and a choice outcome consisting of two or more variables considered simultaneously. It models a worker's day activity pattern—the choice among 162 alternatives for completing the non-work portion of the day—with each alternative defining a configuration of optional tours and commute stops for maintenance and discretionary purposes. The estimated EC is superior to a nested logit model, with significant covariance parameters for several overlapping subsets of the choice set. The case study also demonstrates (a) the importance of testing for simulation error and using many draws in MSL simulation (in this case at least 1500 draws for each simulated probability), and (b) the pending practicality of estimating EC models for large multidimensional choice problems.

LARGE MULTIDIMENSIONAL CHOICE PROBLEMS AND THE ERROR COMPONENT LOGIT KERNEL (EC) MODEL

This paper provides an identification rule and insights for estimation of logit kernel (or mixed logit) models for large multidimensional choice problems. These are problems with large choice sets in which the choice outcome consists of two or more variables considered simultaneously, each variable representing one "dimension" of choice. In the field of transportation the classic large multidimensional problem is the joint mode-destination choice. Current passenger travel choice modeling problems involve many more dimensions, including activity participation—for all purposes—and its accompanying travel modes, destinations and timing for all travel by all household members during the course of a day or even longer periods. In large multidimensional choice problems, it is almost a given that shared unobserved attributes exist among various subsets of alternatives, and heteroscedasticity is also likely.

Model forms exist that capture correlation in multidimensional problems. **Nested logit** (NL) is used for two-dimensional problems in which nests within one dimension share unobserved attributes more significantly than do nests in the other dimension. Ben-Akiva and Lerman *(1)* describe this for the classic mode-destination choice problem; NL must assume that correlation among alternatives sharing the same mode are insignificant, or that correlation among alternatives sharing the same destination are insignificant. NL extends to more dimensions. However, even with two dimensions its conditional independence assumptions can be untenable; ignored correlations may be responsible for difficulties in finding a nested specification with the desired characteristic that all nesting parameters fall between zero and one. The **Paired combinatorial logit** (PCL) *(2)* allows shared unobserved effects in multiple dimensions, defining each pair of alternatives as a dimension. In concept, the approach could probably be generalized to accommodate n-tuples of alternatives, but the problem would become unwieldy, and the generalization has apparently not been attempted. **Cross-nested logit** (CNL) *(3)* generalizes nested logit to accommodate shared unobserved attributes among alternatives within nests in more than one dimension. In the mode-destination problem, it allows for shared unobserved attributes among alternatives with the same mode, and (simultaneously) shared unobserved attributes among alternatives with the same destination. The CNL could probably be generalized to accommodate nesting in three or more dimensions. However, this would also become unwieldy, and has apparently not been implemented. **Heteroscedastic logit** *(4)* accommodates variances that differ across individual alternatives, but it is unable to capture correlation among multiple alternatives in multiple dimensions, a hallmark of multidimensional problems. The above forms all belong to the **Generalized extreme value** (GEV) family of models *(5)* along with several other forms that have been developed to capture a variety of patterns of correlation among alternatives *(6)*.

The **Logit kernel** model used in this paper is more general than, and can be used to approximate any of the above model forms. It is an MNL model with random parameters, and the probabilities are calculated by integrating over the distribution of the random parameters. McFadden and Train *(7)* have proven that it can approximate any random utility model. They call it **Mixed multinomial logit (MMNL)**, because it is essentially a mixture of multiple (infinite if a random parameter has a continuous distribution) multinomial logit models. Train's text *(8)* is an excellent reference regarding MMNL and Bhat *(9)* provides an extensive review of GEV and MMNL models.

The **Factor analytic logit kernel** model, described in detail by Walker *et al (10)* is a particular—but still very flexible—form of logit kernel, based on the factor analytic structure proposed by McFadden *(11)* for the probit model. It takes the following form (using notation of Walker *et al (10)*):

$$U_n = X_n\beta + F_n T\zeta_n + \nu_n \ , \tag{1}$$

$$\text{cov}(U_n) = F_n TT' F_n' + (g/\mu^2) I_{J_n} \tag{2}$$

(denoted as $\Omega_n = \Sigma_n + \Gamma_n$),

where: U_n is a $(J_n \times 1)$ vector of utilities;

X_n is a $(J_n \times K)$ matrix of explanatory variables;

β is a $(K \times 1)$ vector of unknown parameters;

F_n is a $(J_n \times M)$ matrix of factor loadings, including fixed and/or unknown parameters;

T is a $(M \times M)$ lower triangular matrix of unknown parameters, where $TT' = Cov(\xi_n = T\zeta_n)$;

ζ_n is a $(M \times 1)$ vector of i.i.d. random variables with zero mean and unit variance; and

ν_n is a $(J_n \times 1)$ vector of i.i.d. Gumbel random variables with zero location parameter and scale equal to $\mu > 0$. The variance is g/μ^2 , where g is the variance of a standard Gumbel ($\pi^2/6$).

The unknown parameters in this model are μ , β , those in F_n , and those in T . X_n are observed, whereas ζ_n and ν_n are unobserved. The choice probability is $P(i) = \int_\zeta \Lambda(i \mid \zeta) n(\zeta, I_M) d\zeta$, where $n(\zeta, I_M)$ is the joint density function of ζ , and $\Lambda(i \mid \zeta_n) = e^{\mu(X_{in}\beta + F_{in}T\zeta_n)} \Big/ \sum_{j \in C_n} e^{\mu(X_{jn}\beta + F_{jn}T\zeta_n)}$, an MNL model. The probability can be simulated by $\hat{P}(i) = (1/\mathbb{D})\sum_{d=1}^{\mathbb{D}} \Lambda(i \mid \zeta_n^d)$, where ζ_n^d denotes draw d from the distribution of ζ . This simulated probability can be used to simulate the likelihood function for maximum simulated likelihood (MSL) estimation of the parameters. Once the parameters are estimated, it can be used directly (with simulation) for model predictions. In this paper, subsequent references to "**logit kernel**" refer to this factor analytic form of logit kernel.

There are good reasons to use one of the above model forms other than logit kernel if it adequately captures the important correlations. In particular, in each case the choice probability has a closed form, so that estimation and application of the model can be accomplished without resorting to costly numercial approximation techniques such as Gaussian quadrature, simulated likelihood or simulated moments. However, not only can logit kernel mimic all of these forms, it can also combine them with each other and with other forms, such as random parameters, and provides flexibility for extending them without much, if any, further theoretical development. When they are combined and extended, the computational requirements increase, but the basic specification, estimation and application procedures remain the same. Thus, for large multi-dimensional problems, logit kernel is an appealing option.

Table 1 shows the *F* matrix for examples of NL, CNL and PCL analogs, and various combinations. In all cases *F* is composed of zeroes and ones, and the (unshown) *T* matrix is diagonal. Each column of *F* defines a **nest** of alternatives, with a 1 in the column, that share unobserved attributes. Some of these models have nests in multiple dimensions, defining a **dimension** as a group of non-overlapping nests (ie, within a dimension, an alternative can belong to no more than one nest). Nest membership within a dimension is sometimes determined by the value of a categorical attribute, such as mode or destination.

Table 1: Logit kernel models (*F* matrix with shorthand label and description)

<table>
<tr><td>$\begin{bmatrix} 1&0\\1&0\\1&0\\1&0\\0&1\\0&1\\0&1\\0&1 \end{bmatrix}$</td><td>NL(1:2)
Nested logit analog with 2 nests in one dimension</td><td>$\begin{bmatrix} 1&0&0\\0&1&1\\1&1&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$</td><td>PCL(3)
Paired combinatorial logit with 3 pairs</td></tr>
<tr><td>$\begin{bmatrix} 1&0&1&0&0&0\\1&0&1&0&0&0\\1&0&0&1&0&0\\1&0&0&1&0&0\\0&1&0&0&1&0\\0&1&0&0&1&0\\0&1&0&0&0&1\\0&1&0&0&0&1 \end{bmatrix}$</td><td>NL(2:2,4)
2-dimensional nested logit with 2 nests in the first dimension, and 4 nests in the second dimension (2 subnests in each nest of the first dimension)</td><td>$\begin{bmatrix} 1&0&0\\0&1&1\\1&1&0\\1&0&0\\0&1&1\\0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$</td><td>TCL(3)
‘Tripled’ combinatorial logit with 3 triples</td></tr>
<tr><td>$\begin{bmatrix} 1&0&1&0\\1&0&1&0\\1&0&0&1\\1&0&0&1\\0&1&1&0\\0&1&1&0\\0&1&0&1\\0&1&0&1 \end{bmatrix}$</td><td>CNL(2:2,2)
2-dimensional cross-nested logit with 2 nests in the first dimension and 2 nests in the second dimension</td><td>$\left[\begin{array}{cccccc|cc} 1&0&1&0&0&0&1&0\\1&0&1&0&0&0&1&0\\1&0&0&1&0&0&0&1\\1&0&0&1&0&0&0&1\\0&1&0&0&1&0&1&0\\0&1&0&0&1&0&1&0\\0&1&0&0&0&1&0&1\\0&1&0&0&0&1&0&1 \end{array}\right]$</td><td>NL(2:2,4)CNL(1:2)
Combination of NL and CNL</td></tr>
<tr><td>$\begin{bmatrix} 1&0&1&0&1&0&0\\1&0&1&0&1&0&0\\1&0&0&1&1&0&0\\1&0&0&1&0&1&0\\0&1&1&0&0&1&0\\0&1&1&0&0&0&1\\0&1&0&1&0&0&1\\0&1&0&1&0&0&1 \end{bmatrix}$</td><td>CNL(3:2,2,3)
3-dimensional cross-nested logit with 2 nests in each of the first two dimensions, and 3 nests in the third dimension</td><td>$\left[\begin{array}{cc|ccc} 1&0&1&0&0\\1&0&0&1&1\\1&0&1&1&0\\1&0&0&0&0\\0&1&0&0&0\\0&1&0&0&0\\0&1&0&0&0\\0&1&0&0&1 \end{array}\right]$</td><td>NL(1:2)PCL(3)
Combination of NL and PCL</td></tr>
<tr><td></td><td></td><td>$\left[\begin{array}{cccccc|ccccc|ccc} 1&0&1&0&0&0&1&0&1&0&0&1&0&0\\1&0&1&0&0&0&1&0&1&0&0&0&1&1\\1&0&0&1&0&0&0&1&1&0&0&1&1&0\\1&0&0&1&0&0&0&1&0&1&0&0&0&0\\0&1&0&0&1&0&1&0&0&1&0&0&0&0\\0&1&0&0&1&0&1&0&0&0&1&0&0&0\\0&1&0&0&0&1&0&1&0&0&1&0&0&0\\0&1&0&0&0&1&0&1&0&0&1&0&0&1 \end{array}\right]$</td><td>NL(2:2,4)CNL(2:2,3)
PCL(3)
Combination of NL, CNL and PCL</td></tr>
</table>

These models do not address the issue of heteroscedasticity. As specified, including the most general form of *T* matrix (with a distinct parameter on each element of the diagonal), each model imposes a particular form of heteroscedasticity. For example, in NL(1:2), with $T=\begin{bmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{bmatrix}$, the alternatives in the first nest (are assumed to) have equal variance, but different than the variance of alternatives in the second nest.

It is possible to enforce a homoscedasticity assumption or any desired heteroscedasticity assumption by adding columns defining one-alternative nests and specifying a particular form for each diagonal element of *T*. This broadens the class of models to include ALL logit kernel models in which *F* is comprised of ones and zeroes, and T is diagonal, referred to herein as **Error component logit kernel** (EC), as named by Ben-Akiva and Bolduc *(12)*. **Table 2** shows F and T for several EC models, including the analogs of MNL, heteroscedastic logit, and three NL variations. Among the NL variations, only the first replicates a pure NL model, with homoscedastic disturbances and shared unobserved attributes of equal magnitude in both nests.

Table 2: Error component logit kernel models (EC)

F	***T***	**Description**
$\begin{bmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \end{bmatrix}$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0 \\ 0&\sigma_1&0&0&0&0&0&0 \\ 0&0&\sigma_1&0&0&0&0&0 \\ 0&0&0&\sigma_1&0&0&0&0 \\ 0&0&0&0&\sigma_1&0&0&0 \\ 0&0&0&0&0&\sigma_1&0&0 \\ 0&0&0&0&0&0&\sigma_1&0 \\ 0&0&0&0&0&0&0&\sigma_1 \end{bmatrix}$	MNL
$\begin{bmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \end{bmatrix}$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0 \\ 0&\sigma_2&0&0&0&0&0&0 \\ 0&0&\sigma_3&0&0&0&0&0 \\ 0&0&0&\sigma_4&0&0&0&0 \\ 0&0&0&0&\sigma_5&0&0&0 \\ 0&0&0&0&0&\sigma_6&0&0 \\ 0&0&0&0&0&0&\sigma_7&0 \\ 0&0&0&0&0&0&0&\sigma_8 \end{bmatrix}$	Heteroscedastic Logit
$\begin{bmatrix} 1&0 \\ 1&0 \\ 1&0 \\ 1&0 \\ 0&1 \\ 0&1 \\ 0&1 \\ 0&1 \end{bmatrix}$	$\begin{bmatrix} \sigma_1&0 \\ 0&\sigma_1 \end{bmatrix}$	NL(1:2)—Homoscedastic with shared unobserved attributes of same magnitude in both nests
$\left[\begin{array}{cc\|cccccccc} 1&0&1&0&0&0&0&0&0&0 \\ 1&0&0&1&0&0&0&0&0&0 \\ 1&0&0&0&1&0&0&0&0&0 \\ 1&0&0&0&0&1&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0 \\ 0&1&0&0&0&0&0&1&0&0 \\ 0&1&0&0&0&0&0&0&1&0 \\ 0&1&0&0&0&0&0&0&0&1 \end{array}\right]$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0&0&0 \\ 0&\sigma_2&0&0&0&0&0&0&0&0 \\ 0&0&\sigma_2&0&0&0&0&0&0&0 \\ 0&0&0&\sigma_2&0&0&0&0&0&0 \\ 0&0&0&0&\sigma_2&0&0&0&0&0 \\ 0&0&0&0&0&\sigma_2&0&0&0&0 \\ 0&0&0&0&0&0&\sigma_1&0&0&0 \\ 0&0&0&0&0&0&0&\sigma_1&0&0 \\ 0&0&0&0&0&0&0&0&\sigma_1&0 \\ 0&0&0&0&0&0&0&0&0&\sigma_1 \end{bmatrix}$	NL(1:2)—Homoscedastic with shared unobserved attributes of different magnitude in the two nests
$\left[\begin{array}{cc\|cccccccc} 1&0&1&0&0&0&0&0&0&0 \\ 1&0&0&1&0&0&0&0&0&0 \\ 1&0&0&0&1&0&0&0&0&0 \\ 1&0&0&0&0&1&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0 \\ 0&1&0&0&0&0&0&1&0&0 \\ 0&1&0&0&0&0&0&0&1&0 \\ 0&1&0&0&0&0&0&0&0&1 \end{array}\right]$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0&0&0 \\ 0&\sigma_2&0&0&0&0&0&0&0&0 \\ 0&0&\sigma_3&0&0&0&0&0&0&0 \\ 0&0&0&\sigma_4&0&0&0&0&0&0 \\ 0&0&0&0&\sigma_5&0&0&0&0&0 \\ 0&0&0&0&0&\sigma_6&0&0&0&0 \\ 0&0&0&0&0&0&\sigma_7&0&0&0 \\ 0&0&0&0&0&0&0&\sigma_8&0&0 \\ 0&0&0&0&0&0&0&0&\sigma_9&0 \\ 0&0&0&0&0&0&0&0&0&\sigma_{10} \end{bmatrix}$	NL(1:2)—Heteroscedastic with shared unobserved attributes of different magnitude in the two nests

IDENTIFICATION OF EC

Introduction

As Walker *et al (10)* point out, identification of logit kernel is an important and non-trivial task, because of the joint presence of the IID Gumbel disturbance and the probit-like covariance matrix. They review the necessary **Order condition** and sufficient **Rank condition** (for determining the number of identifiable covariance parameters) and what they call the **Equality condition** (for choosing restrictions to uniquely identify the parameters while preserving equality of the restricted differenced covariance matrix with its unrestricted counterpart). They then use these conditions to establish simple sufficient rules for correctly identifying and normalizing two kinds of logit kernel models: one-dimensional nesting (like the simplest nested logit) and heteroscedasticity:

Nesting

$M = 2$ nests	$M - 1$ parameters identified and normalization is arbitrary
$M \geq 3$ nests	M parameters identified

Heteroscedasticity

$J = 2$ alternatives	0 parameters identified
$J \geq 3$ alternatives	$J - 1$ parameters identified and must constrain the minimum variance term to 0

The strengths of the rules are their simplicity and sufficiency. Unfortunately, they only cover a small subset of the EC models with which we are concerned.

This section presents a rule for identification and normalization of all EC models. The rule is necessary but not sufficient; it can be viewed as an enhancement of the order condition, providing a more refined method of discovering identification mistakes, but it does not in all cases guarantee that a particular specification is correctly identified. However, it correctly handles all the particular models presented in this paper. In most cases it identifies and normalizes them. In a few cases it detects the need to rely on a formal (and more tedious) application of the Rank and Equality conditions.

Overview of the EC identification rule

1. **Order condition**. This is the standard Order condition, necessary for identifying any logit kernel model. It provides an upper bound on the number of identifiable parameters that is rarely a binding constraint.
2. **Complementary pairs**. For any complementary pair of nests, at most one covariance parameter can be identified. "**Complementary pair**" means two nests that define two mutually exclusive and collectively exhaustive subsets of the universal choice set. The two nests are called **complements**. In terms of the *F* matrix, a complementary pair consists of two columns of *F* for which there is one '1' and one '0' in each row. Complementary pairs must be restricted in any way that preserves non-negativity for both *T*-elements of the pair and allows their sum of squares to be expressed with a single parameter. (This is satisfied by restricting either member's covariance parameter to zero, or restricting them to equal each other.) This is a generalization of the nesting rule for M=2 nests.
3. **Heteroscedasticity**. At most *J*-1 covariance parameters can be identified for the 2*J* possible columns (*J* complementary pairs) of *F* representing single alternatives and their complements. This broadens Walker, *et al*'s heteroscedasticity rule for $J \geq 3$ alternatives, taking into consideration the complements of the heteroscedastic alternatives. If the columns of *F* include only single alternatives and their complements, then a satisfactory normalization restricts both members of the pair with smallest sum of variance to 0, and employs complementary pair normalization for all other pairs. A simple normalization rule has not been determined for the general class of EC models with a mixture of heteroscedastic and other columns.
4. **Subsets.** Given choice set C of size J, and any subset $\tilde{C}$ of size $\tilde{J}$ = 2, 3, 4 or 5, it may be the case that at most $\sum_{i=1}^{\tilde{J}} i$ parameters (3, 6, 10 or 15, respectively) can be identified for the $2^{\tilde{J}} - 1$ (i.e. 3, 7, 15 or 31, respectively) possible non-empty subsets of $\tilde{C}$ and their $2^{\tilde{J}} - 1$ complements. The pattern suggests that the rule may hold for $\tilde{J}$ >5, but this has not been proven. No simple normalization rule has been determined, although there is evidence that for $\tilde{J}$ =3 the normalization is arbitrary and for $\tilde{J}$ >3 it is not. Fortunately, the subset rule applies only to very complex EC specifications. In many cases, it is possible to simply confirm that the subset rule does not apply, and then to apply the first three parts of the EC identification rule.

The entire rule can be stated succinctly as follows:

EC identification and normalization rule

Identification:

For any EC model (logit kernel with factor matrix *F* comprised of 1s and 0s, and diagonal matrix *T*), no more than

$$I = \min[J(J-1)/2-1, M-C_1+\min(H-C_2, J-1)] \quad (3)$$

covariance parameters can be identified, where:

J is the size of the universal choice set C
M is the number of unique columns of F with at least 2 and less than J-1 nested alternatives
C_1 is the number of complementary pairs among the M columns
H is the number of unique columns with 1 or J-1 nested alternatives
C_2 is the number of complementary pairs among the H columns

Normalization:

All complementary pairs must be restricted in any way that preserves non-negativity for both *T*-elements of the pair and allows their sum of squares to be expressed with a single parameter. (This is satisfied, for example, by restricting either member's *T*-element to zero, or restricting both to equal each other.)

Among the *J* possible complementary pairs of columns with 1 and *J*-1 nested alternatives, an additional normalization may be required. If these are the only columns of *F* , then a satisfactory normalization restricts both members of the pair with smallest sum of variance to 0, and employs complementary pair normalization for all other pairs. A simple normalization rule has not been determined for the general class of EC models with a mixture of heteroscedastic and other columns.

The number of identifiable parameters may be less than *I* if for any choice subset $\tilde{C}$ of size $\tilde{J}$ = 2, 3, 4 or 5, more than $\sum_{i=1}^{\tilde{J}} i$ parameters (3, 6, 10 or 15, respectively) have been specified for the $2^{\tilde{J}}-1$ possible non-empty subsets of $\tilde{C}$ and their $2^{\tilde{J}}-1$ complements.

Support for the EC identification and normalization rule

Complementary pair identification. Walker *et al (10)* demonstrate their nesting rule for *M*=2 nests with the following example (**Example 1**):

$$F=\begin{bmatrix}1&0\\1&0\\0&1\\0&1\\0&1\end{bmatrix},\ T=\begin{bmatrix}\sigma_1&0\\0&\sigma_2\end{bmatrix},\ V(\Omega_\Delta)=\begin{bmatrix}\sigma_{11}+\sigma_{22}+2g/\mu^2\\\sigma_{11}+\sigma_{22}+g/\mu^2\\g/\mu^2\\2g/\mu^2\end{bmatrix},\ J(V)=\begin{bmatrix}1&1&2\\1&1&1\\0&0&1\\0&0&2\end{bmatrix},\ rank(J)-1=1.$$

In addition to the *F* and *T* matrices, the key outputs of a formal test of the rank condition are shown above, including the unique terms of the vectorized differenced covariance matrix, denoted $V(\Omega_\Delta)$, its jacobian, denoted $J(V)$, and the rank of the jacobian minus 1, indicating the number of identifiable parameters . They note that only the sum $(\sigma_{11}+\sigma_{22})$ can be identified. This is because the two variance terms always appear together as a sum in $V(\Omega_\Delta)$, and it results in duplicate columns of $J(V)$. As it turns out, for **every** distinct pair of complementary columns of *F* with a corresponding pair of unique parameters σ_i and σ_j on the diagonal of T, the variance terms σ_{ii} and σ_{jj} always appear together as a sum in $V(\Omega_\Delta)$, and a duplicate pair of columns occurs in $J(V)$. This can be shown to be true regardless of the presence of any other columns of *F* (see **Appendix**), and results in the complementary pair rule.

Example 2 demonstrates the rule in the case where two additional complementary columns are added to the previous example. In this and subsequent examples the *T* matrix is not shown, but is a diagonal matrix with parameter σ_j in column *j*. With the addition of a complementary pair of *F* columns comes a duplicate pair of

$J(V)$ columns, so at most one more parameter is identifiable. In this case no other rules further limit the number of identifiable parameters:

$$F=\begin{bmatrix}1&0&1&0\\1&0&1&0\\0&1&1&0\\0&1&0&1\\0&1&0&1\end{bmatrix},\ V(\Omega_\Delta)=\begin{bmatrix}\sigma_{11}+\sigma_{22}+\sigma_{33}+\sigma_{44}+2g/\mu^2\\\sigma_{11}+\sigma_{22}+\sigma_{33}+\sigma_{44}+g/\mu^2\\\sigma_{33}+\sigma_{44}+g/\mu^2\\g/\mu^2\\\sigma_{33}+\sigma_{44}+2g/\mu^2\\2g/\mu^2\end{bmatrix},\ J(V)=\begin{bmatrix}1&1&1&1&2\\1&1&1&1&1\\0&0&1&1&1\\0&0&0&0&1\\0&0&1&1&2\\0&0&0&0&2\end{bmatrix},\ rank(J)-1=2\,.$$

Complementary pair normalization. Walker *et al (10)* argue that, because the model identifies the sum $(\sigma_{11}+\sigma_{22})$, any normalization of σ_{11} and σ_{22} is acceptable that expresses this sum with one parameter. When a specification includes two or more complementary pairs of nests, the same argument applies independently to each complementary pair.

Heteroscedasticity identification. The complementary pair rule implies that the variance of a lone alternative in a heteroscedastic specification cannot be identified separately from the shared unobserved attributes of the remaining *J*-1 alternatives, because the two columns of *F* corresponding to these two parameters are a complementary pair. This implies that Walker *et al*'s heteroscedasticity rule—that at most *J*-1 parameters can be identified for heteroscedastic terms—can be broadened to say that at most *J*-1 parameters can be identified for the set of all columns of *F* comprising single alternatives and their complements. **Example 3** demonstrates this limit for the case with four alternatives:

$$F=\begin{bmatrix}1&0&0&1&0&1&0&1\\0&1&1&0&0&1&0&1\\0&1&0&1&1&0&0&1\\0&1&0&1&0&1&1&0\end{bmatrix},\ V(\Omega_\Delta)=\begin{bmatrix}\sigma_{11}+\sigma_{22}+\sigma_{77}+\sigma_{88}+2g/\mu^2\\\sigma_{77}+\sigma_{88}+g/\mu^2\\\sigma_{33}+\sigma_{44}+\sigma_{77}+\sigma_{88}+2g/\mu^2\\\sigma_{55}+\sigma_{66}+\sigma_{77}+\sigma_{88}+2g/\mu^2\end{bmatrix},\ J(V)=\begin{bmatrix}1&1&0&0&0&0&1&1&2\\0&0&0&0&0&0&1&1&1\\0&0&1&1&0&0&1&1&2\\0&0&0&0&1&1&1&1&2\end{bmatrix},\ rank(J)-1=3\,.$$

Heteroscedasticity normalization. The broadened identification rule makes it apparent that in a pure heteroscedastic specification, an implicit normalization is occuring; the parameter of each alternative's complement is assumed to equal zero. If this assumption is relaxed for any of these complements, the normalization rule for complementary pairs allows for any normalization of the pair that expresses the sum with one parameter. In such a case, however, Walker *et al*'s normalization rule must broaden to deal with the complements. It is satisfactory to set to zero the smallest sum $(\sigma_{ii}+\sigma_{jj})$ among the complementary pairs involving a heteroscedastic alternative. Walker *et al's* normalization argument still applies, with $(\sigma_{ii}+\sigma_{jj})$ substituting for the heteroscedastic variance. In essence, σ_{jj} was lurking in their argument, but with an assumed value of 0.

However, a simple normalization rule has not been determined for the general class of EC models with a mixture of heteroscedastic (including complements) **and** other columns, where it is possible that correlations among alternatives might cause the Equality condition to be violated under the simple heteroscedasticity normalization.

Subsets. As it turns out, the order condition, complementary pair rule and heteroscedasticy rule, when taken together, as in equation (3) are necessary (as argued above) but not sufficient. Additional limits apply when many nests are specified for any particular subset of alternatives. Given a model with a choice set C of size J , any subset $\tilde{C}$ of size $\tilde{J}$ has $2^{\tilde{J}}-1$ possible nests (including nests of one alternative), and for $\tilde{J}$ >2, it may be that less than $2^{\tilde{J}}-1$ of them can be identified. Empirical tests of the rank condition have in all cases yielded results consistent with the following table of identification limits:

Table 3: Identification limits for subsets of EC nests

$\tilde{J}$	maximum number of nests within any subset $\tilde{C}$ of size $\tilde{J}$ ($2^{\tilde{J}}-1$)	maximum number of identifiable nest parameters within any subset $\tilde{C}$ of size $\tilde{J}$ ($\sum_{i=1}^{\tilde{J}} i$)
2	3	3
3	7	6
4	15	10
5	31	15

Example 4 illustrates a case where J=8 , and the top 3 ($\tilde{J}$) rows of F correspond to $\tilde{C}$:

$$F=\begin{bmatrix}1110100\\1101010\\1011001\\0000000\\0000000\\0000000\\0000000\\0000000\end{bmatrix}, V(\Omega_\Delta)=\begin{bmatrix}\sigma_{11}+\sigma_{22}+\sigma_{33}+\sigma_{55}+2g/\mu^2\\ \sigma_{11}+\sigma_{22}+g/\mu^2\\ \sigma_{11}+\sigma_{33}+g/\mu^2\\ g/\mu^2\\ \sigma_{11}+\sigma_{22}+\sigma_{44}+\sigma_{66}+2g/\mu^2\\ \sigma_{11}+\sigma_{44}+g/\mu^2\\ \sigma_{11}+\sigma_{33}+\sigma_{44}+\sigma_{77}+2g/\mu^2\\ 2g/\mu^2\end{bmatrix}, \quad J(V)=\begin{bmatrix}11101002\\11000001\\10100001\\00000001\\11010102\\10010001\\10110012\\00000002\end{bmatrix}, \quad rank(J)-1=6 .$$

Seven nests have been defined for $\tilde{C}$, but only 6 parameters are identified; in $J(V)$, columns 1, 5, 6 and 7 are a linear combination of columns 2, 3 and 4.

Using the EC identification and normalization rule

The rule can be translated into the following practical procedure for specifying, identifying and normalizing an EC model:

EC identification and normalization procedure

Step 1: Hypothesize a non-redundant set of nests. It may be helpful, although not necessary, to first specify sets of nests that span the choice set and then remove those that are undesired or redundant.

Step 2: Count the M nests with more than 1 and less than J-1 alternatives, find the C_1 complementary pairs among them, and remove one member of each pair.

Step 3: Count the H nests with 1 or J-1 alternatives (heteroscedastic nests), find the C_2 complementary pairs among them, and remove one member of each pair.

Step 4: If there is a full set of heteroscedastic nests (H-C_2=J), then one more restriction of the heteroscedastic nests is required, although it is not clear at this point what restriction(s) will not violate the Equality condition. Even if there is not a full set of heteroscedastic nests, it is possible that the heteroscedastic nests violate the Equality condition. In either case, it may be necessary to estimate the model with more than one alternative normalization, to try to empirically determine the unbiased normalization.

Step 5: Verify that the order condition is not violated—that the remaining number of covariance parameters is less than $J(J$-1)/2.

Step 6: For each covariance parameter in the model, list the nests in the model that are subsets of it or its complement. Remove one or more subset nests for each case where it appears that the number of identifiable subsets is exceeded.

CASE STUDY: THE ACTIVITY PATTERN MODEL

The activity pattern model

The case study involves the modeling of secondary activity and travel choices of workers on their workday. Together with the choice of the day's primary activity purpose and location (in this case, work on tour), this model defines the day's activity pattern *(13)*. The activity pattern model is a major integrating component of the activity schedule model system, which also includes conditional models of locations, timing and travel mode for all the activity pattern's stops. Even by itself, and conditioned by primary activity, the activity pattern model is large and multidimensional. In this case study it has a choice set of 162 alternatives defined by the feasible combinations of outcomes (shown in parentheses below) in four primary dimensions:

- Presence and purpose of stops on the way to work (none, maintenance, discretionary)
- Presence and purpose of work-based subtours (none, maintenance, discretionary)
- Presence and purpose of stops on the way home from work (none, maintenance, discretionary)
- Number and purpose of home-based tours (0, 1 maintenance, 2+ maintenance, 1 discretionary, 2+ discretionary, 1+ maintenance with 1+ discretionary)

The case study starts with a nested logit model estimated from 1994 Portland Metro diary survey data, and tests for the presence of shared unobserved attributes in several dimensions.

Hypothesized EC parameters and their identification

We use the six step procedure presented above to hypothesize, identify and normalize the EC parameters of the model:

Step 1: **Table 4** lists ten dimensions, each spanning the choice set, with suspected shared unobserved attributes. In hypothesizing the importance of dimension 1, for example, we suspect that the lone pattern alternative with no extra stops or tours is viewed as distinctly different from the other 161 pattern alternatives in ways that the utility functions cannot capture. In dimension 2, we similarly suspect an otherwise unexplained preference for, or against, the six pattern alternatives with no extra stops on the commute tour.

The first nest in dimensions 1, 9 and 10 is redundant, so we remove it from dimensions 9 and 10.

Table 4: Hypothesized EC nests of the activity pattern model

Dim-en-sion	Kept Nest	Description	No. of alts
1	1	Pattern has no secondary tours or stops	1
		Pattern has secondary tours and/or stops	161
2	2	Commute tour has extra stops	156
		Commute tour has no extra stops	6
3	3	Pattern has secondary tours	135
		Pattern has no secondary tours	27
4	4	Pattern has secondary maintenance tours and/or maintenance stops on primary tour	138
		Pattern has no maintenance tours and no maintenance stops on primary tour	24
5	5	Pattern has secondary discretionary tours and/or discretionary stops on primary tour	138
		Pattern has no discretionary tours and no discretionary stops on primary tour	24
6	6	Commute has stops before work	108
		Commute has no stops before work	54
7	7	Commute has stops after work	108
		Commute has no stops after work	54
8	8	Pattern has work-based subtour	108
		Pattern has no work-based subtour	54
9		Pattern has no secondary tours or stops	1
	9	Pattern has secondary tours but no secondary stops on primary tour	5
	10	Pattern has secondary stops on primary tour but no secondary tours	26
	11	Pattern has secondary tours and secondary stops on primary tour	130
10		Pattern has no secondary tours or stops	1
	12	Pattern has maintenance tours and/or stops but not discretionary	23
	13	Pattern has discretionary tours and/or stops but not maintenance	23
	14	Pattern has maintenance and discretionary tours and/or stops	115

Step 2: Of the 20 (M) nests with more than 1 and less than J-1 alternatives, the 14 in dimensions 2 through 8 form 7 (C_2) complementary pairs. So, we remove one nest from each of dimensions 2 through 8.

Step 3: Only dimension 1 is left with nests having 1 or J-1 alternatives (ie, H=2), and it is a complementary pair (C_1=1). So, according to the normalization rule, we must remove one of the pair, and it doesn't matter which one. In this case the rule implies that it is impossible to distinguish between variance in the preference for a simple pattern with no secondary tours or stops, on the one hand, and shared unobserved attributes among all the other alternatives, on the other hand. Empirically, they have identical effects. In other words, an unusual preference for (or against) the simplest pattern is empirically no different than an unusual preference against (or for) all of the other alternatives arising from unobserved attribute(s) that they share.

Step 4: Of the 14 remaining nests—numbered as Kept Nests in Table 4—only nest 1 has 1 or J-1 alternatives, so an additional restriction is therefore not necessary. But it is still possible that this normalization of the 162 possible heteroscedastic nests could violate the Equality condition. In the absence of nests 2 through 14, this would only occur if the specified heteroscedastic alternative happened to have smaller variance than all the other alternatives. But in the presence of nests 2-14, the evaluation of the chosen normalization is less straightforward. In either case, we are left to assume that the normalization of the heteroscedastic alternative is okay, or to try and evaluate it empirically.

Step 5: With 162 alternatives and only 14 kept nests, the Order condition is easily satisfied.

Step 6: In 12 cases, a nest in the specification, or its complement, has subsets that are also in the specification (or a subset's complement is in the specification.) These cases are noted in **Table 5**, where each row presents a nest, its size, and its subsets:

Table 5: Cases where identification restrictions may be needed because of EC nest subsets

Case	Nest ($\tilde{C}$)	$\tilde{J}$	subsets of $\tilde{C}$
1	1-complement	161	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
2	2	156	6, 7, 8, 10, 11
3	2-complement	6	1, 9
4	3	135	9, 11
5	3-complement	27	1, 10
6	4	138	12, 14
7	4-complement	24	1, 13
8	5	138	13, 14
9	5-complement	24	1, 12
10	6-complement	54	1, 2-complement
11	7-complement	54	1, 2-complement
12	8-complement	54	1, 2-complement

For example, in Case 3 the complement of nest 2 (patterns in which the commute tour has no extra stops) consists of 6 alternatives, and it has two subset nests in the specification, including nests 1 (no secondary stops or tours) and 9 (secondary tours, but no secondary stops on primary tour). All of the nests with subsets have more than 5 alternatives, so the subset rule does not provide an explicit upper limit on the number of identifiable subsets. However, each case has less than the 15 subset limit for nests with 5 alternatives, so it is unlikely that the presence of subsets restricts the number of identifiable parameters.

Since the identification rule is necessary, but not sufficient, in order to be certain of identification and proper normalization we would need to complete the daunting task of explicitly verifying the rank and equality conditions. However, the just-completed analysis gives us a high degree of confidence that this EC model specification is identified.

Model estimation

We specified and estimated two EC specifications. The first, EC1, is the EC analog of the original NL model, including only the dimension 1 covariance parameter. The second, EC14, includes all 14 covariance parameters, and is not augmented to maintain the NL homoscedasticity assumption. Both models were estimated using the Maximum Simulated Likelihood (MSL) feature of ALOGIT 4EC *(14)*, with probabilities simulated from

Shuffled Halton draws *(15)*. Shuffled Halton draws differ from a Halton sequence in one respect: the sequence is randomly shuffled before drawing. The Shuffled Halton method was chosen because, like the pure Halton or scrambled Halton methods, it covers the simulation space more efficiently than quasi-random draws but, unlike them, it does not suffer from high inter-dimensional correlation when the specification has a large number of dimensions. Estimation using shuffled uniform vectors *(16)* is another simple and even more efficient method that could have been used. This is like shuffled Halton, except it shuffles evenly spaced numbers instead of a Halton sequence.

Table 6 shows the estimation results for the two EC models, as well as maximum likelhood estimation results for the original MNL and NL models. It also includes an additional EC model (EC4), in which 10 insignificant EC14 covariance parameters are eliminated. All five models include the same 57 utility function parameters. Parameters 1-15 associate utility with each maintenance and discretionary activity in a pattern, regardless of the activity's position in the pattern. Parameters 16-55 alter a pattern alternative's utility, depending on the position of maintenance and discretionary activities in the pattern. These utility effects of the presence and position of maintenance and dicretionary activities depend heavily on the characteristics of the individual and the individual's household. Parameters 56 and 57 capture the sensitivity of pattern choice to land development and transportation attributes (such as activity opportunities, mode availability, travel time and travel cost), via expected utility measures (logsums) from conditional destination, mode, and time-of-day choice models. For a detailed discussion of the factors affecting pattern choice, and the integration of activity pattern models with travel demand models, see Bowman *(17)*.

One important note regarding parameters 1-57 is that their values are not perfectly comparable across the models. As more structural parameters (58-71) are introduced, more of the variation in the data is attributed to structural correlations, so less is attributed to the IID disturbance of the MNL component of the model, increasing its scale. Since the IID scale is embedded in the utility function parameters, 1-57, its increase causes across the board increases in their estimated values.

Our primary interest is in the covariance parameters, 58-71, where the potential advantage of the EC specification over the MNL and NL models can be tested. Looking first at parameter 58, we note that the NL and NL analog (EC1) both capture the distinctiveness of the pattern with no secondary tours or stops (or, viewed another way, perceived but unspecified similarity of the other 161 alternatives). The effect is statistically quite significant. The sign and magnitude of the parameters is different, but this is because they measure the same covariance in different ways. The same effect is even more pronounced in EC14, evidenced in the larger magnitude of parameter 58 in this model. Looking at the rest of the EC14 parameters, we see that, although the NL specifications apparently capture the strongest nesting effect, several other parameters are also large and significant, especially 68, 64, and 70. These capture the effect of shared unobserved attributes among three additional distinct (but overlapping) subsets of alternatives: the 130 patterns with secondary tours AND secondary stops on the commute tour (68), the 108 patterns with stops on the way home from work (64), and the 23 patterns with discretionary stops and/or tours but not maintenance (70). An additional model (EC4) was estimated, retaining only these four covariance parameters. The loglikelihood values are not perfectly comparable, because the EC models have simulated likelihood functions. However, it appears that moving from NL to EC4 obtains nearly as much likelihood improvement as moving from MNL to NL. The EC4 adjusted rho squared is also the best of the tested models, by a slight margin. These results, coupled with the intuitive appeal and very strong significance of the four EC4 covariance parameters, point to its superiority. However, model application was not conducted as part of this research, and the statistical results are close enough between NL and EC4 that a comparison of model application results would seem appropriate before finally choosing one specification over the other. If EC4 performed essentially the same as NL in relevant forecasting tasks, then its extra operating overhead costs might not be justified. On the other hand, reverting to NL would ignore the possibility that the EC4 might perform differently than NL in an important way in some yet untested forecasting task.

Table 6: Estimation results for work-on-tour activity pattern model with 162 alternatives

			MNL		NL		EC1 (NL analog)		EC14		EC4	
	Pattern component	**Variable description**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**
	Utility Function Parameters											
1	Secondary on-tour maintenance activities	constant	-2.725	0.25	-2.820	0.26	-2.819	0.26	-3.089	0.34	-3.030	0.33
2		child age 12-17	-0.529	0.30	-0.462	0.35	-0.452	0.34	-0.632	0.48	-0.536	0.41
3		#children 0-17, age 18+ male	0.064	0.03	0.055	0.03	0.054	0.03	0.090	0.04	0.073	0.04
4		#children 0-17, age 18+ female	0.219	0.03	0.244	0.03	0.241	0.03	0.336	0.05	0.301	0.04
5		child age 18+	-0.401	0.13	-0.408	0.13	-0.407	0.13	-0.558	0.18	-0.483	0.15
6		per capita income ($10K)	0.061	0.02	0.075	0.02	0.074	0.02	0.094	0.03	0.083	0.03
7		workforce participation rate	0.237	0.10	0.288	0.12	0.291	0.11	0.390	0.16	0.336	0.14
8	Secondary on-tour discretionary activities	constant	-2.866	0.24	-2.820	0.24	-2.819	0.24	-3.146	0.31	-3.108	0.31
9		child age 12-17	0.271	0.30	0.302	0.35	0.311	0.34	0.312	0.46	0.279	0.42
10		nonfamily, age 18+	0.127	0.08	0.136	0.09	0.136	0.09	0.173	0.11	0.177	0.10
11		children 0-11 are in HH, age 18+ female	-0.456	0.12	-0.451	0.12	-0.454	0.12	-0.503	0.15	-0.456	0.13
12		full-time worker	-0.260	0.10	-0.356	0.11	-0.356	0.11	-0.357	0.13	-0.380	0.12
13		student	-0.234	0.10	-0.250	0.10	-0.249	0.10	-0.311	0.12	-0.276	0.11
14		per capita income, full time worker	0.050	0.02	0.061	0.03	0.061	0.03	0.070	0.03	0.066	0.03
15		children 0-11 are in HH, age 18+ male	-0.136	0.08	-0.164	0.08	-0.164	0.08	-0.239	0.11	-0.202	0.10
16	**Position of secondary on-tour maintenance activity in pattern, base case is after primary activity** maintenance stop before primary activity	constant	-0.249	0.14	-0.212	0.14	-0.213	0.14	-0.008	0.21	0.049	0.21
17	maintenance dest-based subtour	constant	-0.110	0.23	-0.049	0.23	-0.050	0.23	0.101	0.30	0.155	0.29
18		children 0-11 are in HH, age 18+ female	-0.817	0.20	-0.857	0.20	-0.856	0.20	-0.959	0.20	-0.939	0.20
19	Secondary maintenance tour	Constant	1.165	0.33	1.310	0.34	1.312	0.34	1.582	0.42	1.551	0.41
20		fulltime worker	-0.527	0.09	-0.598	0.10	-0.599	0.10	-0.633	0.11	-0.629	0.11
21		per capita income	-0.137	0.03	-0.139	0.03	-0.139	0.03	-0.165	0.04	-0.158	0.04
22		1+ cars per adult	-0.291	0.08	-0.311	0.08	-0.309	0.08	-0.349	0.10	-0.322	0.09
23	**Position of secondary on-tour discretionary activity in pattern, base case is after primary activity** before primary activity	Constant	-0.521	0.16	-0.487	0.16	-0.487	0.16	-0.207	0.23	-0.187	0.23
24	discretionary dest-based subtour	Constant	0.311	0.28	0.356	0.28	0.355	0.28	0.575	0.34	0.607	0.34
25		fulltime worker	0.956	0.17	0.976	0.17	0.976	0.17	0.996	0.18	0.989	0.17
26	Secondary discretionary tour	Constant	0.292	0.41	0.342	0.42	0.345	0.42	0.569	0.51	0.528	0.49
27		family w children 0-11, age 18+ female	0.273	0.16	0.316	0.16	0.317	0.16	0.410	0.17	0.386	0.16
28		per capita income	-0.158	0.03	-0.159	0.03	-0.159	0.03	-0.186	0.04	-0.178	0.04
29		1+ cars per adult	0.283	0.10	0.275	0.10	0.276	0.10	0.311	0.12	0.297	0.11

Table 6: Estimation results for work-on-tour activity pattern model with 162 alternatives

			MNL		NL		EC1 (NL analog)		EC14		EC4	
	Pattern component	**Variable description**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**
30	**Secondary activity combinations on primary tour** Maintenance stops before & after	constant	1.055	0.12	1.012	0.12	1.012	0.12	0.890	0.13	0.877	0.13
31		children 0-4 are in household, age 18+	0.566	0.21	0.600	0.22	0.598	0.22	0.660	0.23	0.666	0.23
32		children 0-4 are in household, age 18+ female	0.616	0.26	0.557	0.27	0.561	0.27	0.531	0.29	0.540	0.29
33	No secondary stops on primary tour	children age 5-11 are in HH, age 18+	-0.190	0.10	-0.204	0.10	-0.205	0.10	-0.300	0.12	-0.281	0.11
34		self-employed, age 18+	-0.236	0.15	-0.267	0.15	-0.267	0.15	-0.361	0.18	-0.340	0.16
35		age 35-49	-0.121	0.07	-0.128	0.07	-0.128	0.07	-0.145	0.09	-0.134	0.08
36		age 18+, female	-0.182	0.06	-0.267	0.07	-0.267	0.07	-0.329	0.09	-0.306	0.08
37	**Inter-tour combinations, simple primary tour with 0 or 1 secondary tours and complex primary tour with 0 secondary tours are base cases** simple primary tour with 2+ secondary tours	constant	-1.646	0.14	-1.774	0.15	-1.774	0.15	-1.899	0.18	-1.882	0.15
38	complex primary tour with 1 secondary tour	constant	-0.630	0.08	-0.733	0.09	-0.733	0.09	-2.202	0.57	-1.906	0.54
39	complex primary tour with 2+ secondary tours	constant	-2.553	0.20	-2.799	0.21	-2.799	0.21	-4.365	0.64	-4.090	0.60
40	2+ secondary tours	children age 12-17 are in HH, age 18+	0.490	0.15	0.503	0.15	0.504	0.15	0.439	0.16	0.459	0.16
41	simple primary tour with 0 secondary tours	fulltime worker	0.239	0.08	0.159	0.12	0.161	0.12	0.354	0.17	0.251	0.14
42		children age 5-11 are in HH, age 18+	-0.081	0.11	-0.233	0.15	-0.231	0.15	-0.284	0.20	-0.224	0.18
43		HH income over $60K	-0.307	0.07	-0.478	0.13	-0.471	0.13	-0.631	0.18	-0.574	0.16
44		child age 18+	0.015	0.15	0.195	0.23	0.199	0.23	0.347	0.32	0.303	0.29
45		nonfamily, age 18+	0.172	0.09	0.176	0.12	0.175	0.12	0.332	0.17	0.297	0.15
46		self-employed, age 18+	-0.140	0.16	-0.315	0.23	-0.303	0.22	-0.449	0.29	-0.390	0.26
47		student, age 18+	-0.460	0.12	-0.694	0.20	-0.688	0.20	-0.929	0.28	-0.815	0.24
48		no vehicles in HH, age 18+	0.355	0.17	0.507	0.26	0.514	0.27	0.709	0.37	0.665	0.33
49		age 35-49	-0.044	0.08	-0.161	0.11	-0.162	0.11	-0.226	0.15	-0.204	0.14
50		HH income is under $30K, age 18+	0.244	0.08	0.261	0.11	0.264	0.11	0.383	0.16	0.335	0.14
51		children age 0-4 are in HH, age 18+	0.292	0.09	0.310	0.10	0.311	0.10	0.406	0.13	0.378	0.11
52	**Secondary stop and tour purposes included in pattern** maintenance and discretionary	children age 5-11 are in HH, age 18+	0.168	0.10	0.164	0.10	0.164	0.10	0.350	0.13	0.316	0.12
53		2+ adults in HH, age 18+	-0.150	0.07	-0.146	0.07	-0.147	0.07	-0.127	0.11	-0.123	0.08
54		age 18-24	-0.367	0.13	-0.347	0.14	-0.348	0.14	-0.429	0.19	-0.369	0.15
55		nonfamily, age 18+	0.229	0.10	0.281	0.10	0.282	0.10	0.315	0.15	0.236	0.12

Table 6: Estimation results for work-on-tour activity pattern model with 162 alternatives

			MNL		NL		EC1 (NL analog)		EC14		EC4	
	Pattern component	**Variable description**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**
56	**Logsums from the nested mode-dest-TOD choice models of the pattern's tours** Secondary tours	weighted avg of 4 sec. tour types' mode-dest-TOD logsums. For patterns with 2+ sec. tours, logsum is scaled up by avg no. of sec. tours observed in sample for pattern type and tour purpose.	0.445	0.09	0.469	0.09	0.468	0.09	0.504	0.11	0.515	0.10
57	Primary work tour		1.358	0.32	1.458	0.32	1.457	0.32	1.283	0.37	1.328	0.37
58	**Covariance Parameters**	Pattern has no sec. tours or stops			0.644	0.09	-1.874	0.45	-2.773	0.62	-2.468	0.56
59		Commute tour has extra stops							-0.465	0.83		
60		Pattern has sec. tours							-0.059	0.46		
61		Pattern has sec. maint tours and/or maint stops on primary tour							-1.050	0.50		
62		Pattern has sec. discret tours and/or discret stops on primary tour							0.015	1.10		
63		Commute has stops before work							-0.205	0.44		
64		Commute has stops after work							-1.669	0.48	-1.605	0.46
65		Pattern has work-based subtour							-0.128	0.39		
66		Pattern has sec. tours but no sec. stops on primary tour							-0.301	0.61		
67		Pattern has sec. stops on primary tour but no sec. tours							-0.028	0.63		
68		Pattern has sec. tours and sec. stops on primary tour							-2.548	0.63	-2.173	0.64
69		Pattern has maint tours and/or stops but not discret							-0.572	0.89		
70		Pattern has discret tours and/or stops but not maint							-1.270	0.31	-1.079	0.24
71		Pattern has maint and discret tours and/or stops							-0.495	0.34		
	Estimation Method Summary											
	Method			ML		NL		MSL		MSL		MSL
	Type of draws for MSL							Shuff. Halton		Shuff. Halton		Shuff. Halton
	No. of draws per simulated probability							100		1500		1500
	Summary Statistics											
	Number observed choices			6170		6170		6170		6170		6170
	Number of estimated parameters			57		58		58		71		61
	Log likelihood w coeffs=0			-31390		-31390		-31390		-31390		-31390
	Final Log likelihood			-16890		-16885		-16885		-16876		-16880
	Rho squared			0.46193		0.46211		0.46209		0.46238		0.46225
	Adjusted rho squared			0.46011		0.46026		0.46025		0.46012		0.46030

Using EC for large multidimensional models

The undertaken empirical work provides some insight regarding important questions about the development of large EC models. The **first question** is, how many Shuffled Halton draws are necessary? The answer is to use enough draws so that simulation error does not significantly affect the estimation results. Researchers who have considered this issue in the past have used small models and re-estimated many times, enabling them to empirically measure the variability of the MSL estimation results. Unfortunately, with the EC14 model, the time required for estimation practically prevents the use of this approach. Nevertheless, we can gain some insight by estimating the model a small number of times, varying the number of draws used for the simulated probabilities. EC14 was estimated with 100, 400 and 1500 draws per simulated probability, using the convergent estimates from the 100 draw case as starting values for the 400 and 1500 draw cases. **Table 7** shows the estimated EC14 covariance parameters, with the 4 EC4 parameters highlighted. Although parameter 61 is significant in the 1500 draw case, it was insignificant in a subsequent re-estimation with 1500 draws using a new seed, and again in an EC5 model (EC4 with parameter 61 added). The EC4 parameters are of similar magnitude and significance with 1500 and 400 draws, but two of them are insignificant in the 100 draw case; if only 100 draws had been used, the selected model might have been "EC4", but with a different set of four covariance parameters. This provides the first empirical clue that 100 draws is insufficient.

Table 7: EC14 covariance parameter estimates with 100, 400, and 1500 draws per simulated probability

	100 draws			400 draws			1500 draws		
	est.	std error	abs T	est.	std error	abs T	est.	std error	abs T
58	**-3.054**	**0.70**	**4.34**	**-2.882**	**0.62**	**4.63**	**-2.773**	**0.62**	**4.45**
59	0.957	0.47	2.02	-0.097	0.50	0.19	-0.465	0.83	0.56
60	0.881	0.48	1.85	-0.143	0.39	0.36	-0.059	0.46	0.13
61	0.100	0.43	0.23	-0.105	0.41	0.26	-1.050	0.50	2.08
62	-0.078	0.68	0.11	-0.529	0.58	0.91	0.015	1.10	0.01
63	-0.127	0.33	0.39	-0.609	0.33	1.85	-0.205	0.44	0.46
64	**-0.032**	**0.26**	**0.13**	**1.878**	**0.47**	**4.00**	**-1.669**	**0.48**	**3.48**
65	0.051	0.28	0.18	-0.216	0.31	0.70	-0.128	0.39	0.33
66	0.013	0.28	0.05	0.878	0.61	1.44	-0.301	0.61	0.49
67	-0.846	0.45	1.86	-0.040	0.42	0.10	-0.028	0.63	0.05
68	**-0.164**	**0.68**	**0.24**	**-2.552**	**0.58**	**4.40**	**-2.548**	**0.63**	**4.03**
69	-1.436	0.39	3.72	-0.683	0.43	1.60	-0.572	0.89	0.65
70	**-0.951**	**0.27**	**3.54**	**-1.234**	**0.29**	**4.21**	**-1.270**	**0.31**	**4.13**
71	-0.143	0.29	0.49	0.620	0.27	2.26	-0.495	0.34	1.48

In a further test for simulation error, the EC4 model was estimated two times with 100 draws per simulated probability, using the same starting values but different simulation seeds, and statistics were calculated measuring the magnitude of the difference in the results. This procedure was repeated with 400 and 1500 draws, respectively, and the results are shown in **Table 8.** The original intent was to repeat this comparison with 5000 draws, but that turned out to be prohibitively time consuming.

The statistics in the row numbered 1 show that the root mean squared (RMS) percentage change in the covariance parameter estimates is almost 25% when comparing two 100 draw simulations, and drops to 10.6% with 400 draws, and to 3.5% with 1500 draws. Rows 2 through 4 show that the maximum percent change in covariance parameters, as well as the root mean squared and maximum changes in absolute T, have drops of similar magnitude. The rows numbered 5 through 8 provide the same summary measures for the utility function parameters. Here the difference between root mean squared and maximum changes is more pronounced, probably because several utility function parameters have small absolute T values, but the increase from 100 to 1500 draws again shows a 5 to 10-fold reduction in change to well under 10%. Since each of the statistics is based on only a pair of MSL estimation runs, it is possible that luck has caused the 1500 simulation case to appear superior to the 100 and 400 simulation cases. So, while we can't be sure that 1500 simulations is adequate without re-estimating more times, these results show fairly large variation with only 100 simulations.

Table 8: Summary of changes in EC4 Maximum Simulated Likelihood (MSL) estimation results when the Shuffled Halton simulation seed changes

	Summary Statistic	Two 100 draw simulations	Two 400 draw simulations	Two 1500 draw simulations
A. Excluding cases with |T|<1				
	Covariance Parameters (all statistics based on absolute values of parameter estimates)			
	Percentage of cases with |T|<1	25.0%	0.0%	0.0%
1	Root mean squared percent change in parameter estimates	23.3%	10.6%	3.5%
2	Maximum percent change in parameter estimates	39.1%	18.2%	5.9%
3	Root mean squared percent change in |T|	16.7%	15.7%	3.5%
4	Maximum percent change in |T|	27.8%	27.5%	6.1%
	Utility Function Parameters			
	Percentage of cases with |T|<1	7.0%	8.8%	7.0%
5	Root mean squared percent change in parameter estimates	11.7%	4.4%	1.4%
6	Maximum percent change in parameter estimates	57.9%	20.4%	5.6%
7	Root mean squared percent change in |T|	11.7%	4.1%	1.3%
8	Maximum percent change in |T|	65.0%	18.9%	4.5%
B. Including cases with |T|<1				
	Covariance Parameters (all statistics based on absolute values of parameter estimates)			
9	Root mean squared percent change in parameter estimates	4844.0%	10.6%	3.5%
10	Maximum percent change in parameter estimates	9688.0%	18.2%	5.9%
11	Root mean squared percent change in |T|	6489.1%	15.7%	3.5%
12	Maximum percent change in |T|	12978.2%	27.5%	6.1%
	Utility Function Parameters			
13	Root mean squared percent change in parameter estimates	23.4%	57.6%	3.7%
14	Maximum percent change in parameter estimates	125.8%	431.4%	23.8%
15	Root mean squared percent change in |T|	22.4%	57.0%	3.5%
16	Maximum percent change in |T|	123.6%	427.6%	21.7%

Rows 1 through 8 exclude parameters with absolute T values smaller than 1, because they could arguably be removed from the model and they tend to dominate the change statistics. Rows 9 through 16 repeat the statistics, this time including the parameters with small absolute T. In the 400 and 1500 simulation cases, there is no change in the covariance parameter statistics, because none of the parameters has absolute T less than 1. For the utility parameters, the presence of a few with small absolute T causes the statistics to get somewhat larger. In the case of 100 draws, the covariance parameter statistics explode. This occurs because the lone insignificant covariance parameter became quite large and significant in the second 100 simulation estimation run, and it underscores the insufficiency of the 100 draw simulation.

The above analysis suggests that **for models of this size, estimated by MSL with Shuffled Halton draws, too much simulation error may occur when using only 100 draws per simulated probability, and using 1500 draws or more is desirable. It is advisable to estimate the model at least twice (with different simulation seeds), using the selected number of draws per simulated probability, in order to gain a sense of the variability caused by the simulation procedure**.

The **second question** is, how long does it take to estimate large multidimensional EC models? Does it take so long that it is still impractical, in most cases, to estimate them (ignoring the application run times for now)? **Table 9** shows that the EC14 specification with 1500 draws per probability calculation took 22000 times longer than the MNL per iteration, and 80000 times longer overall, because it required more iterations. All the models in the table have 162 alternatives, 6170 observed choices, 57 utility parameters, and were estimated using ALOGIT 4EC. All models except the largest were estimated on a dedicated IBM Thinkpad T30 laptop computer with a 1.8 GHz Intel Pentium 4 processor and 256megabytes of memory, operating under the Windows 2000 operating system. The largest model was estimated on a desktop machine with a 3.0 GHz Intel Pentium 4 processor and 1gigabyte of memory, operating under Windows XP Pro. In order to compare its runtimes to those estimated on the laptop, the EC model was estimated on both platforms with 100 and 400 draws per simulated probability, to establish a conversion factor of .57 on estimation run times. On the more powerful machine, the total run time was 28 days, too long to be practical except in the research lab. However, the numerical aspects of this model estimation problem are well-suited to a programming implementation in which multiple processors solve the problem efficiently by working in parallel. It is easy to imagine a 20 CPU array of processors, each operating 5 times as fast as the 3MHz machine, solving this problem in 6 hours, which would be quite practical. So, **at this time, it is not very practical to estimate EC models for large multidimensional problems, but that should change within the next few years as computational power increases, or sooner if estimation software becomes available that supports parallel processing**.

Table 9: Model estimation run time comparisons. All the models in the table have 162 alternatives, 6170 observed choices, 57 utility parameters, and were estimated using ALOGIT 4EC. Except where noted, run times are for a dedicated IBM Thinkpad T30 laptop computer with a 1.8 GHz Intel Pentium 4 processor and 256megabytes of memory, operating under the Windows 2000 operating system.

Model	MNL	NL	EC1	EC14	EC14	EC14
Number of covariance parameters	0	1	1	14	14	14
Number of draws per simulated probability			100	100	400	1500
Previously converged model supplying starting values:		MNL	NL	EC1	EC14 (100 draws)	EC14 (100 draws)
Number of iterations to convergence	9	8	5	14	25	24
iteration duration (seconds)	7.6	25.8	410	11250	45000	170526
Approximate iteration duration	8 seconds	26 seconds	7 minutes	3 hours	12 hours	2 days
iteration duration (relative to MNL)		3x	50x	1500x	6000x	22000x
Approximate total estimation duration	1 minute	5 minutes	30 minutes	2 days	13 days	50 days
total duration (relative to MNL)		5x	30x	3000x	20000x	80000x
Estimation duration on Windows XPPro machine with 3 MHz Pentium 4 processor						28 days

The table notes that convergent values of simpler models were used as starting values for the more complex models. Although it makes the comparison of run times difficult, it was done with the hope of avoiding numerical problems in the optimization search routine and reducing the number of required iterations. Numerical problems were avoided, but the only case where it seems clear that iterations were reduced is in the estimation of the true NL model using convergent MNL results as starting values. Using the convergent EC14 100 draw results might have increased the required iterations of the 400 and 1500 draw versions; the differences in estimated parameters discussed above indicates that the search routine was forced to traverse fairly large changes in estimated values in order to converge on new MSL estimates. Therefore, **it may save time to start MSL estimation of EC models using the anticipated acceptable number of draws, rather than starting with few draws and increasing the number in subsequent estimation runs.**

CONCLUSION

The identification rule, with accompanying procedure, presented in this paper provides a manageable method of increasing confidence that a large error component logit kernel (EC) model is identified, and the case study demonstrates its use. The case study also demonstrates the statistical advantages of logit kernel for a particular large multidimensional choice problem, identifies some development pitfalls to avoid, and concludes that using EC with maximum simulated likelihood should soon be practical for large real-world choice modeling problems.

ACKNOWLEDGMENTS

The author recognizes the important work of Walker, *et al* (10) on identification, upon which this paper builds. Joan Walker was especially helpful in reviewing a preliminary draft, discussing the topic, and suggesting enhancements. Moshe Ben-Akiva, Denis Bolduc, Dinesh Gopinath and others gave valuable feedback during a preliminary presentation of the research results at an MIT seminar. Kenneth Train and four anonymous referees provided very helpful written suggestions. The author accepts sole responsibility for any remaining errors.

APPENDIX—PROOF OF COMPLEMENTARY PAIR IDENTIFICATION RULE

Given an EC model with *F* matrix (JxM) having a complementary pair of columns *a* and *b*, and *T* matrix (MxM) having elements $T_{aa} = \sigma_a$ and $T_{bb} = \sigma_b$, and no other appearances in *T* of σ_a or σ_b, then σ_{aa} and σ_{bb} always appear together as a sum in the vectorized differenced covariance matrix $V(\Omega_\Delta)$, and a duplicate pair of columns occurs in its jacobian, $J(V)$.

Proof: Let Δ be the (J-1)xJ differencing matrix, consisting of the (J-1)x(J-1) identity matrix augmented by a negative unit column vector. Define the four sets of indices, $\boldsymbol{A} \equiv \{i \mid F_{ia} = 1\}$, $\boldsymbol{A^c} \equiv \{i \mid F_{ia} = 0\}$, $\boldsymbol{B} \equiv \{i \mid F_{ib} = 1\}$, $\boldsymbol{B^c} \equiv \{i \mid F_{ib} = 0\}$. Then, since *a* and *b* are a complementary pair, $\boldsymbol{A} = \boldsymbol{B^c}$ and $\boldsymbol{A^c} = \boldsymbol{B}$. It follows that in the covariance matrix $\Omega \equiv FTT'F'$, any element Ω_{ij} includes the additive term σ_{aa} exactly once if and only if $i \in \boldsymbol{A}$ and $j \in \boldsymbol{A}$, and it does not include any function of σ_a otherwise. Likewise, any element $\Omega_{ij} \in \Omega$ includes the additive term σ_{bb} exactly once if and only if $i \in \boldsymbol{A^c}$ and $j \in \boldsymbol{A^c}$, and it does not include any function of σ_b otherwise. Therefore, Ω includes four types of elements: (1) $\Omega_{ij} \mid i \in \boldsymbol{A}, j \in \boldsymbol{A}$, containing the additive term σ_{aa} exactly once and no function of σ_b; (2) $\Omega_{ij} \mid i \in \boldsymbol{A}, j \in \boldsymbol{A^c}$, containing no function of either σ_a or σ_b; (3) $\Omega_{ij} \mid i \in \boldsymbol{A^c}, j \in \boldsymbol{A}$, containing no function of either σ_a or σ_b; and (4) $\Omega_{ij} \mid i \in \boldsymbol{A^c}, j \in \boldsymbol{A^c}$, containing the additive term σ_{bb} exactly once and no function of σ_a. Each element $\Omega_{\Delta ij}$ of the differenced covariance matrix $\Omega_\Delta \equiv \Delta\Omega\Delta'$ therefore includes the additive term $\sigma_{aa} + \sigma_{bb}$ exactly once (with no other appearance of functions of either σ_a or σ_b) or not at all. It follows directly that σ_{aa} and σ_{bb} always appear together as a sum in $V(\Omega_\Delta)$, and a duplicate pair of columns occurs in $J(V)$. *QED.* This proof can be easily visualized by carrying out its calculations on a representative case.

REFERENCES

1. Ben-Akiva, M. and S. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, Cambridge, Massachusetts, 1985.
2. Chu, C. A Paired Combiinatorial Logit Model for Travel Demand Analysis. In *Proceedings of the Fifth World Conference on Transportation Research, Vol. 4*, Ventura, CA, 1989, pp. 295-309.
3. Vovsha, P. Application of Cross-Nested Logit Model to Mode Choice in Tel-Aviv, Israel, Metropolitan Area. In *Transportation Research Record 1607*, TRB, National Research Council, Washington, D.C., 1997, pp.6-15.
4. Bhat, C. A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice. *Transportation Research B*, Vol. 29, No. 6, 1995, pp. 471-483.
5. McFadden, D. Modeling the choice of residential location. In *Spatial Interaction Theory and Planning Models*, A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, eds., North-Holland, Amsterdam, 1978, pp. 75–96. Available at http://emlab.berkeley.edu/users/mcfadden/dlmcv10.html.
6. Koppelman, F. and V. Sethi. Closed Form Discrete Choice Models. In *Handbook of Transport Modeling*, D.A. Hensher and K.J. Button, eds., Pergamon Press, Oxford, 2000. Available at http://www.civil.northwestern.edu/trans/abstractsfskres.htm#discrete.
7. McFadden, D. and K. Train. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, vol. 15, issue 5, 2000, pp. 447-470. Available at http://emlab.berkeley.edu/users/mcfadden/dlmcv10.html.

8. Train, K. *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, UK, 2003. Available at http://elsa.berkeley.edu/~train/distant.html.
9. Bhat, C. Econometric Models of Choice: Formulation and Estimation. Presented at 10th International Conference on Travel Behaviour Research, Lucerne, August, 2003. Available at http://www.ivt.baum.ethz.ch/allgemein/iatbr2003.html.
10. Walker, J.. M. Ben-Akiva and D. Bolduc. Identification of the Logit Kernel (or Mixed Logit) Model. 2003. Working paper, available at http://mit.edu/jwalker/www/home.htm
11. McFadden, D. Econometric Analysis of Qualitative Response Models, in *Handbook of Econometrics II*, Z. Griliches and M.D. Intriligator, Eds., Elsevier Science Publishers, 1984, pp. 1396-1456.
12. Ben-Akiva, M. and D. Bolduc. Multinomial Probit with a Logit Kernel and a General Parametric Specification of the Covariance Structure. 1996. Working paper.
13. Bowman, J.L., and M. Ben-Akiva. Activity-based disaggregate travel demand model system with activity schedules, *Transportation Research Part A*, 35, 2001, pp. 1-28. Available at http://jbowman.net.
14. *ALOGIT 4EC*. HCG Holding. Software documentation available at http://www.hpgholding.nl/software/welcsoft.htm
15. Hess, S. and J. W. Polak. The Shuffled Halton sequence. 2003. Submitted to *Mathematical and Computer modelling*, and available as working paper at http://www.cts.cv.ic.ac.uk/html/Staff/staffDetails.asp?id=SH.
16. Hess, S., K. Train, and J. W. Polak. Use of randomly shifted and shuffled uniform vectors in the estimation of a Mixed Logit model for vehicle choice. 2003. Submitted to Transportation Research Part B and available as working paper at http://www.cts.cv.ic.ac.uk/html/Staff/staffDetails.asp?id=SH.
17. Bowman, J.L. *The Day Activity Schedule Approach to Travel Demand Analysis*, Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, 1998. Available at http://jbowman.net.

List of Tables and Figures

Table 1: Logit kernel models (*F* matrix with shorthand label and description)

$\begin{bmatrix} 1&0\\1&0\\1&0\\1&0\\0&1\\0&1\\0&1\\0&1 \end{bmatrix}$	NL(1:2) Nested logit analog with 2 nests in one dimension	$\begin{bmatrix} 1&0&0\\0&1&1\\1&1&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$	PCL(3) Paired combinatorial logit with 3 pairs
$\begin{bmatrix} 1&0&1&0&0&0\\1&0&1&0&0&0\\1&0&0&1&0&0\\1&0&0&1&0&0\\0&1&0&0&1&0\\0&1&0&0&1&0\\0&1&0&0&0&1\\0&1&0&0&0&1 \end{bmatrix}$	NL(2:2,4) 2-dimensional nested logit with 2 nests in the first dimension, and 4 nests in the second dimension (2 subnests in each nest of the first dimension)	$\begin{bmatrix} 1&0&0\\0&1&1\\1&1&0\\1&0&0\\0&1&1\\0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$	TCL(3) 'Tripled' combinatorial logit with 3 triples
$\begin{bmatrix} 1&0&1&0\\1&0&1&0\\1&0&0&1\\1&0&0&1\\0&1&1&0\\0&1&1&0\\0&1&0&1\\0&1&0&1 \end{bmatrix}$	CNL(2:2,2) 2-dimensional cross-nested logit with 2 nests in the first dimension and 2 nests in the second dimension	$\left[\begin{array}{cccccc\|cc} 1&0&1&0&0&0&1&0\\1&0&1&0&0&0&1&0\\1&0&0&1&0&0&0&1\\1&0&0&1&0&0&0&1\\0&1&0&0&1&0&1&0\\0&1&0&0&1&0&1&0\\0&1&0&0&0&1&0&1\\0&1&0&0&0&1&0&1 \end{array}\right]$	NL(2:2,4)CNL(1:2) Combination of NL and CNL
$\begin{bmatrix} 1&0&1&0&1&0&0\\1&0&1&0&1&0&0\\1&0&0&1&1&0&0\\1&0&0&1&0&1&0\\0&1&1&0&0&1&0\\0&1&1&0&0&0&1\\0&1&0&1&0&0&1\\0&1&0&1&0&0&1 \end{bmatrix}$	CNL(3:2,2,3) 3-dimensional cross-nested logit with 2 nests in each of the first two dimensions, and 3 nests in the third dimension	$\left[\begin{array}{cc\|ccc} 1&0&1&0&0\\1&0&0&1&1\\1&0&1&1&0\\1&0&0&0&0\\0&1&0&0&0\\0&1&0&0&0\\0&1&0&0&0\\0&1&0&0&1 \end{array}\right]$	NL(1:2)PCL(3) Combination of NL and PCL
		$\left[\begin{array}{cccccc\|ccccc\|ccc} 1&0&1&0&0&0&1&0&1&0&0&1&0&0\\1&0&1&0&0&0&1&0&1&0&0&0&1&1\\1&0&0&1&0&0&0&1&1&0&0&1&1&0\\1&0&0&1&0&0&0&1&0&1&0&0&0&0\\0&1&0&0&1&0&1&0&0&1&0&0&0&0\\0&1&0&0&1&0&1&0&0&0&1&0&0&0\\0&1&0&0&0&1&0&1&0&0&1&0&0&0\\0&1&0&0&0&1&0&1&0&0&1&0&0&1 \end{array}\right]$	NL(2:2,4)CNL(2:2,3) PCL(3) Combination of NL, CNL and PCL

Table 2: Error component logit kernel models (EC)

F	***T***	**Description**
$\begin{bmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \end{bmatrix}$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0 \\ 0&\sigma_1&0&0&0&0&0&0 \\ 0&0&\sigma_1&0&0&0&0&0 \\ 0&0&0&\sigma_1&0&0&0&0 \\ 0&0&0&0&\sigma_1&0&0&0 \\ 0&0&0&0&0&\sigma_1&0&0 \\ 0&0&0&0&0&0&\sigma_1&0 \\ 0&0&0&0&0&0&0&\sigma_1 \end{bmatrix}$	MNL
$\begin{bmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \end{bmatrix}$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0 \\ 0&\sigma_2&0&0&0&0&0&0 \\ 0&0&\sigma_3&0&0&0&0&0 \\ 0&0&0&\sigma_4&0&0&0&0 \\ 0&0&0&0&\sigma_5&0&0&0 \\ 0&0&0&0&0&\sigma_6&0&0 \\ 0&0&0&0&0&0&\sigma_7&0 \\ 0&0&0&0&0&0&0&\sigma_8 \end{bmatrix}$	Heteroscedastic Logit
$\begin{bmatrix} 1&0 \\ 1&0 \\ 1&0 \\ 1&0 \\ 0&1 \\ 0&1 \\ 0&1 \\ 0&1 \end{bmatrix}$	$\begin{bmatrix} \sigma_1&0 \\ 0&\sigma_1 \end{bmatrix}$	NL(1:2)—Homoscedastic with shared unobserved attributes of same magnitude in both nests
$\left[\begin{array}{cc\|cccccccc} 1&0&1&0&0&0&0&0&0&0 \\ 1&0&0&1&0&0&0&0&0&0 \\ 1&0&0&0&1&0&0&0&0&0 \\ 1&0&0&0&0&1&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0 \\ 0&1&0&0&0&0&0&1&0&0 \\ 0&1&0&0&0&0&0&0&1&0 \\ 0&1&0&0&0&0&0&0&0&1 \end{array}\right]$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0&0&0 \\ 0&\sigma_2&0&0&0&0&0&0&0&0 \\ 0&0&\sigma_2&0&0&0&0&0&0&0 \\ 0&0&0&\sigma_2&0&0&0&0&0&0 \\ 0&0&0&0&\sigma_2&0&0&0&0&0 \\ 0&0&0&0&0&\sigma_2&0&0&0&0 \\ 0&0&0&0&0&0&\sigma_1&0&0&0 \\ 0&0&0&0&0&0&0&\sigma_1&0&0 \\ 0&0&0&0&0&0&0&0&\sigma_1&0 \\ 0&0&0&0&0&0&0&0&0&\sigma_1 \end{bmatrix}$	NL(1:2)—Homoscedastic with shared unobserved attributes of different magnitude in the two nests
$\left[\begin{array}{cc\|cccccccc} 1&0&1&0&0&0&0&0&0&0 \\ 1&0&0&1&0&0&0&0&0&0 \\ 1&0&0&0&1&0&0&0&0&0 \\ 1&0&0&0&0&1&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0 \\ 0&1&0&0&0&0&0&1&0&0 \\ 0&1&0&0&0&0&0&0&1&0 \\ 0&1&0&0&0&0&0&0&0&1 \end{array}\right]$	$\begin{bmatrix} \sigma_1&0&0&0&0&0&0&0&0&0 \\ 0&\sigma_2&0&0&0&0&0&0&0&0 \\ 0&0&\sigma_3&0&0&0&0&0&0&0 \\ 0&0&0&\sigma_4&0&0&0&0&0&0 \\ 0&0&0&0&\sigma_5&0&0&0&0&0 \\ 0&0&0&0&0&\sigma_6&0&0&0&0 \\ 0&0&0&0&0&0&\sigma_7&0&0&0 \\ 0&0&0&0&0&0&0&\sigma_8&0&0 \\ 0&0&0&0&0&0&0&0&\sigma_9&0 \\ 0&0&0&0&0&0&0&0&0&\sigma_{10} \end{bmatrix}$	NL(1:2)—Heteroscedastic with shared unobserved attributes of different magnitude in the two nests

Table 3: Identification limits for subsets of EC nests

$\tilde{J}$	maximum number of nests within any subset $\tilde{C}$ of size $\tilde{J}$ ($2^{\tilde{J}} - 1$)	maximum number of identifiable nest parameters within any subset $\tilde{C}$ of size $\tilde{J}$ ($\sum_{i=1}^{\tilde{J}} i$)
2	3	3
3	7	6
4	15	10
5	31	15

Table 4: Hypothesized EC nests of the activity pattern model

Dim-en-sion	Kept Nest	Description	No. of alts
1	1	Pattern has no secondary tours or stops	1
		Pattern has secondary tours and/or stops	161
2	2	Commute tour has extra stops	156
		Commute tour has no extra stops	6
3	3	Pattern has secondary tours	135
		Pattern has no secondary tours	27
4	4	Pattern has secondary maintenance tours and/or maintenance stops on primary tour	138
		Pattern has no maintenance tours and no maintenance stops on primary tour	24
5	5	Pattern has secondary discretionary tours and/or discretionary stops on primary tour	138
		Pattern has no discretionary tours and no discretionary stops on primary tour	24
6	6	Commute has stops before work	108
		Commute has no stops before work	54
7	7	Commute has stops after work	108
		Commute has no stops after work	54
8	8	Pattern has work-based subtour	108
		Pattern has no work-based subtour	54
9		Pattern has no secondary tours or stops	1
	9	Pattern has secondary tours but no secondary stops on primary tour	5
	10	Pattern has secondary stops on primary tour but no secondary tours	26
	11	Pattern has secondary tours and secondary stops on primary tour	130
10		Pattern has no secondary tours or stops	1
	12	Pattern has maintenance tours and/or stops but not discretionary	23
	13	Pattern has discretionary tours and/or stops but not maintenance	23
	14	Pattern has maintenance and discretionary tours and/or stops	115

Table 5: Cases where identification restrictions may be needed because of EC nest subsets

Case	Nest ($\tilde{C}$)	$\tilde{J}$	subsets of $\tilde{C}$
1	1-complement	161	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
2	2	156	6, 7, 8, 10, 11
3	2-complement	6	1, 9
4	3	135	9, 11
5	3-complement	27	1, 10
6	4	138	12, 14
7	4-complement	24	1, 13
8	5	138	13, 14
9	5-complement	24	1, 12
10	6-complement	54	1, 2-complement
11	7-complement	54	1, 2-complement
12	8-complement	54	1, 2-complement

Table 6: Estimation results for work-on-tour activity pattern model with 162 alternatives

			MNL		NL		EC1 (NL analog)		EC14		EC4	
	Pattern component	**Variable description**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**
	Utility Function Parameters											
1	Secondary on-tour maintenance activities	constant	-2.725	0.25	-2.820	0.26	-2.819	0.26	-3.089	0.34	-3.030	0.33
2		child age 12-17	-0.529	0.30	-0.462	0.35	-0.452	0.34	-0.632	0.48	-0.536	0.41
3		#children 0-17, age 18+ male	0.064	0.03	0.055	0.03	0.054	0.03	0.090	0.04	0.073	0.04
4		#children 0-17, age 18+ female	0.219	0.03	0.244	0.03	0.241	0.03	0.336	0.05	0.301	0.04
5		child age 18+	-0.401	0.13	-0.408	0.13	-0.407	0.13	-0.558	0.18	-0.483	0.15
6		per capita income ($10K)	0.061	0.02	0.075	0.02	0.074	0.02	0.094	0.03	0.083	0.03
7		workforce participation rate	0.237	0.10	0.288	0.12	0.291	0.11	0.390	0.16	0.336	0.14
8	Secondary on-tour discretionary activities	constant	-2.866	0.24	-2.820	0.24	-2.819	0.24	-3.146	0.31	-3.108	0.31
9		child age 12-17	0.271	0.30	0.302	0.35	0.311	0.34	0.312	0.46	0.279	0.42
10		nonfamily, age 18+	0.127	0.08	0.136	0.09	0.136	0.09	0.173	0.11	0.177	0.10
11		children 0-11 are in HH, age 18+ female	-0.456	0.12	-0.451	0.12	-0.454	0.12	-0.503	0.15	-0.456	0.13
12		full-time worker	-0.260	0.10	-0.356	0.11	-0.356	0.11	-0.357	0.13	-0.380	0.12
13		student	-0.234	0.10	-0.250	0.10	-0.249	0.10	-0.311	0.12	-0.276	0.11
14		per capita income, full time worker	0.050	0.02	0.061	0.03	0.061	0.03	0.070	0.03	0.066	0.03
15		children 0-11 are in HH, age 18+ male	-0.136	0.08	-0.164	0.08	-0.164	0.08	-0.239	0.11	-0.202	0.10
16	**Position of secondary on-tour maintenance activity in pattern, base case is after primary activity** maintenance stop before primary activity	constant	-0.249	0.14	-0.212	0.14	-0.213	0.14	-0.008	0.21	0.049	0.21
17	maintenance dest-based subtour	constant	-0.110	0.23	-0.049	0.23	-0.050	0.23	0.101	0.30	0.155	0.29
18		children 0-11 are in HH, age 18+ female	-0.817	0.20	-0.857	0.20	-0.856	0.20	-0.959	0.20	-0.939	0.20
19	Secondary maintenance tour	Constant	1.165	0.33	1.310	0.34	1.312	0.34	1.582	0.42	1.551	0.41
20		fulltime worker	-0.527	0.09	-0.598	0.10	-0.599	0.10	-0.633	0.11	-0.629	0.11
21		per capita income	-0.137	0.03	-0.139	0.03	-0.139	0.03	-0.165	0.04	-0.158	0.04
22		1+ cars per adult	-0.291	0.08	-0.311	0.08	-0.309	0.08	-0.349	0.10	-0.322	0.09
23	**Position of secondary on-tour discretionary activity in pattern, base case is after primary activity** before primary activity	Constant	-0.521	0.16	-0.487	0.16	-0.487	0.16	-0.207	0.23	-0.187	0.23
24	discretionary dest-based subtour	Constant	0.311	0.28	0.356	0.28	0.355	0.28	0.575	0.34	0.607	0.34
25		fulltime worker	0.956	0.17	0.976	0.17	0.976	0.17	0.996	0.18	0.989	0.17
26	Secondary discretionary tour	Constant	0.292	0.41	0.342	0.42	0.345	0.42	0.569	0.51	0.528	0.49
27		family w children 0-11, age 18+ female	0.273	0.16	0.316	0.16	0.317	0.16	0.410	0.17	0.386	0.16
28		per capita income	-0.158	0.03	-0.159	0.03	-0.159	0.03	-0.186	0.04	-0.178	0.04
29		1+ cars per adult	0.283	0.10	0.275	0.10	0.276	0.10	0.311	0.12	0.297	0.11

Table 6: Estimation results for work-on-tour activity pattern model with 162 alternatives

			MNL		NL		EC1 (NL analog)		EC14		EC4	
	Pattern component	**Variable description**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**
30	**Secondary activity combinations on primary tour** Maintenance stops before & after	constant	1.055	0.12	1.012	0.12	1.012	0.12	0.890	0.13	0.877	0.13
31		children 0-4 are in household, age 18+	0.566	0.21	0.600	0.22	0.598	0.22	0.660	0.23	0.666	0.23
32		children 0-4 are in household, age 18+ female	0.616	0.26	0.557	0.27	0.561	0.27	0.531	0.29	0.540	0.29
33	No secondary stops on primary tour	children age 5-11 are in HH, age 18+	-0.190	0.10	-0.204	0.10	-0.205	0.10	-0.300	0.12	-0.281	0.11
34		self-employed, age 18+	-0.236	0.15	-0.267	0.15	-0.267	0.15	-0.361	0.18	-0.340	0.16
35		age 35-49	-0.121	0.07	-0.128	0.07	-0.128	0.07	-0.145	0.09	-0.134	0.08
36		age 18+, female	-0.182	0.06	-0.267	0.07	-0.267	0.07	-0.329	0.09	-0.306	0.08
37	**Inter-tour combinations, simple primary tour with 0 or 1 secondary tours and complex primary tour with 0 secondary tours are base cases** simple primary tour with 2+ secondary tours	constant	-1.646	0.14	-1.774	0.15	-1.774	0.15	-1.899	0.18	-1.882	0.15
38	complex primary tour with 1 secondary tour	constant	-0.630	0.08	-0.733	0.09	-0.733	0.09	-2.202	0.57	-1.906	0.54
39	complex primary tour with 2+ secondary tours	constant	-2.553	0.20	-2.799	0.21	-2.799	0.21	-4.365	0.64	-4.090	0.60
40	2+ secondary tours	children age 12-17 are in HH, age 18+	0.490	0.15	0.503	0.15	0.504	0.15	0.439	0.16	0.459	0.16
41	simple primary tour with 0 secondary tours	fulltime worker	0.239	0.08	0.159	0.12	0.161	0.12	0.354	0.17	0.251	0.14
42		children age 5-11 are in HH, age 18+	-0.081	0.11	-0.233	0.15	-0.231	0.15	-0.284	0.20	-0.224	0.18
43		HH income over $60K	-0.307	0.07	-0.478	0.13	-0.471	0.13	-0.631	0.18	-0.574	0.16
44		child age 18+	0.015	0.15	0.195	0.23	0.199	0.23	0.347	0.32	0.303	0.29
45		nonfamily, age 18+	0.172	0.09	0.176	0.12	0.175	0.12	0.332	0.17	0.297	0.15
46		self-employed, age 18+	-0.140	0.16	-0.315	0.23	-0.303	0.22	-0.449	0.29	-0.390	0.26
47		student, age 18+	-0.460	0.12	-0.694	0.20	-0.688	0.20	-0.929	0.28	-0.815	0.24
48		no vehicles in HH, age 18+	0.355	0.17	0.507	0.26	0.514	0.27	0.709	0.37	0.665	0.33
49		age 35-49	-0.044	0.08	-0.161	0.11	-0.162	0.11	-0.226	0.15	-0.204	0.14
50		HH income is under $30K, age 18+	0.244	0.08	0.261	0.11	0.264	0.11	0.383	0.16	0.335	0.14
51		children age 0-4 are in HH, age 18+	0.292	0.09	0.310	0.10	0.311	0.10	0.406	0.13	0.378	0.11
52	**Secondary stop and tour purposes included in pattern** maintenance and discretionary	children age 5-11 are in HH, age 18+	0.168	0.10	0.164	0.10	0.164	0.10	0.350	0.13	0.316	0.12
53		2+ adults in HH, age 18+	-0.150	0.07	-0.146	0.07	-0.147	0.07	-0.127	0.11	-0.123	0.08
54		age 18-24	-0.367	0.13	-0.347	0.14	-0.348	0.14	-0.429	0.19	-0.369	0.15
55		nonfamily, age 18+	0.229	0.10	0.281	0.10	0.282	0.10	0.315	0.15	0.236	0.12

Table 6: Estimation results for work-on-tour activity pattern model with 162 alternatives

			MNL		NL		EC1 (NL analog)		EC14		EC4	
	Pattern component	**Variable description**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**	**est.**	**st. err**
56	**Logsums from the nested mode-dest-TOD choice models of the pattern's tours** Secondary tours	weighted avg of 4 sec. tour types' mode-dest-TOD logsums. For patterns with 2+ sec. tours, logsum is scaled up by avg no. of sec. tours observed in sample for pattern type and tour purpose.	0.445	0.09	0.469	0.09	0.468	0.09	0.504	0.11	0.515	0.10
57	Primary work tour		1.358	0.32	1.458	0.32	1.457	0.32	1.283	0.37	1.328	0.37
58	**Covariance Parameters**	Pattern has no sec. tours or stops			0.644	0.09	-1.874	0.45	-2.773	0.62	-2.468	0.56
59		Commute tour has extra stops							-0.465	0.83		
60		Pattern has sec. tours							-0.059	0.46		
61		Pattern has sec. maint tours and/or maint stops on primary tour							-1.050	0.50		
62		Pattern has sec. discret tours and/or discret stops on primary tour							0.015	1.10		
63		Commute has stops before work							-0.205	0.44		
64		Commute has stops after work							-1.669	0.48	-1.605	0.46
65		Pattern has work-based subtour							-0.128	0.39		
66		Pattern has sec. tours but no sec. stops on primary tour							-0.301	0.61		
67		Pattern has sec. stops on primary tour but no sec. tours							-0.028	0.63		
68		Pattern has sec. tours and sec. stops on primary tour							-2.548	0.63	-2.173	0.64
69		Pattern has maint tours and/or stops but not discret							-0.572	0.89		
70		Pattern has discret tours and/or stops but not maint							-1.270	0.31	-1.079	0.24
71		Pattern has maint and discret tours and/or stops							-0.495	0.34		
	Estimation Method Summary											
	Method			ML		NL		MSL		MSL		MSL
	Type of draws for MSL							Shuff. Halton		Shuff. Halton		Shuff. Halton
	No. of draws per simulated probability							100		1500		1500
	Summary Statistics											
	Number observed choices			6170		6170		6170		6170		6170
	Number of estimated parameters			57		58		58		71		61
	Log likelihood w coeffs=0			-31390		-31390		-31390		-31390		-31390
	Final Log likelihood			-16890		-16885		-16885		-16876		-16880
	Rho squared			0.46193		0.46211		0.46209		0.46238		0.46225
	Adjusted rho squared			0.46011		0.46026		0.46025		0.46012		0.46030

Table 7: EC14 covariance parameter estimates with 100, 400, and 1500 draws per simulated probability

	100 draws			400 draws			1500 draws		
	est.	std error	abs T	est.	std error	abs T	est.	std error	abs T
58	**-3.054**	**0.70**	**4.34**	**-2.882**	**0.62**	**4.63**	**-2.773**	**0.62**	**4.45**
59	0.957	0.47	2.02	-0.097	0.50	0.19	-0.465	0.83	0.56
60	0.881	0.48	1.85	-0.143	0.39	0.36	-0.059	0.46	0.13
61	0.100	0.43	0.23	-0.105	0.41	0.26	-1.050	0.50	2.08
62	-0.078	0.68	0.11	-0.529	0.58	0.91	0.015	1.10	0.01
63	-0.127	0.33	0.39	-0.609	0.33	1.85	-0.205	0.44	0.46
64	**-0.032**	**0.26**	**0.13**	**1.878**	**0.47**	**4.00**	**-1.669**	**0.48**	**3.48**
65	0.051	0.28	0.18	-0.216	0.31	0.70	-0.128	0.39	0.33
66	0.013	0.28	0.05	0.878	0.61	1.44	-0.301	0.61	0.49
67	-0.846	0.45	1.86	-0.040	0.42	0.10	-0.028	0.63	0.05
68	**-0.164**	**0.68**	**0.24**	**-2.552**	**0.58**	**4.40**	**-2.548**	**0.63**	**4.03**
69	-1.436	0.39	3.72	-0.683	0.43	1.60	-0.572	0.89	0.65
70	**-0.951**	**0.27**	**3.54**	**-1.234**	**0.29**	**4.21**	**-1.270**	**0.31**	**4.13**
71	-0.143	0.29	0.49	0.620	0.27	2.26	-0.495	0.34	1.48

Table 8: Summary of changes in EC4 Maximum Simulated Likelihood (MSL) estimation results when the Shuffled Halton simulation seed changes

	Summary Statistic	Two 100 draw simulations	Two 400 draw simulations	Two 1500 draw simulations
A. Excluding cases with |T|<1				
	Covariance Parameters (all statistics based on absolute values of parameter estimates)			
	Percentage of cases with |T|<1	25.0%	0.0%	0.0%
1	Root mean squared percent change in parameter estimates	23.3%	10.6%	3.5%
2	Maximum percent change in parameter estimates	39.1%	18.2%	5.9%
3	Root mean squared percent change in |T|	16.7%	15.7%	3.5%
4	Maximum percent change in |T|	27.8%	27.5%	6.1%
	Utility Function Parameters			
	Percentage of cases with |T|<1	7.0%	8.8%	7.0%
5	Root mean squared percent change in parameter estimates	11.7%	4.4%	1.4%
6	Maximum percent change in parameter estimates	57.9%	20.4%	5.6%
7	Root mean squared percent change in |T|	11.7%	4.1%	1.3%
8	Maximum percent change in |T|	65.0%	18.9%	4.5%
B. Including cases with |T|<1				
	Covariance Parameters (all statistics based on absolute values of parameter estimates)			
9	Root mean squared percent change in parameter estimates	4844.0%	10.6%	3.5%
10	Maximum percent change in parameter estimates	9688.0%	18.2%	5.9%
11	Root mean squared percent change in |T|	6489.1%	15.7%	3.5%
12	Maximum percent change in |T|	12978.2%	27.5%	6.1%
	Utility Function Parameters			
13	Root mean squared percent change in parameter estimates	23.4%	57.6%	3.7%
14	Maximum percent change in parameter estimates	125.8%	431.4%	23.8%
15	Root mean squared percent change in |T|	22.4%	57.0%	3.5%
16	Maximum percent change in |T|	123.6%	427.6%	21.7%

Table 9: Model estimation run time comparisons. All the models in the table have 162 alternatives, 6170 observed choices, 57 utility parameters, and were estimated using ALOGIT 4EC. Except where noted, run times are for a dedicated IBM Thinkpad T30 laptop computer with a 1.8 GHz Intel Pentium 4 processor and 256megabytes of memory, operating under the Windows 2000 operating system.

Model	MNL	NL	EC1	EC14	EC14	EC14
Number of covariance parameters	0	1	1	14	14	14
Number of draws per simulated probability			100	100	400	1500
Previously converged model supplying starting values:		MNL	NL	EC1	EC14 (100 draws)	EC14 (100 draws)
Number of iterations to convergence	9	8	5	14	25	24
iteration duration (seconds)	7.6	25.8	410	11250	45000	170526
Approximate iteration duration	8 seconds	26 seconds	7 minutes	3 hours	12 hours	2 days
iteration duration (relative to MNL)		3x	50x	1500x	6000x	22000x
Approximate total estimation duration	1 minute	5 minutes	30 minutes	2 days	13 days	50 days
total duration (relative to MNL)		5x	30x	3000x	20000x	80000x
Estimation duration on Windows XPPro machine with 3 MHz Pentium 4 processor						28 days