A PAIRED COMBINATORIAL LOGIT MODEL FOR TRAVEL DEMAND ANALYSIS

Chaushie Chu

Schimpeler•Corradino Associates
433 S. Spring Street, Suite 1004N
Los Angeles, California 90013

1. INTRODUCTION

In the last decade, the multinomial logit (MNL) model has undergone an immense amount of theoretical and empirical developments (1, 2). Nowadays MNL models are being estimated and applied as a major mode choice component in numerous urban transportation planning studies (3, 4, 5, 6). Despite its wide acceptance, the MNL model has a limitation in accounting for differential substitutability among alternatives, particularly as embodied in its independence of irrelevant alternatives (IIA) property. This has led to an interest in developing alternative model structures that do not possess the IIA property.

Multinomial probit (MNP, 7) and cross-correlated logit (CCL, 8) are the two most flexible model structures that fully relax the IIA property of the MNL model. Using these structures, the alternatives in the choice set can be assumed with any degree of correlation. However, both of these models are mathematically intractable. Numerical integration, which is often time consuming, is usually required when the choice probability is to be computed. Daganzo applied Clark's approximation to save computational effort for MNP models; however, when the number of alternatives in the choice set exceeds 4 or 5, the choice probabilities produced by the approximate method are inaccurate. With the present speed of computers, MNP and CCL models are still limited to academic research contexts.

Other model structures simpler than MNP and CCL are available. Universal logit (UL) nested logit (NL), and ordered logit (OL) are the popular ones. UL was developed by McFadden (9) in the early 1970's. This model is inconsistent with the paradigm of random utility maximization (RUM). But, it has been used frequently as a diagnostic tool for detecting the IIA property in MNL models. NL and OL models were developed in the late 1970's and early 1980's by McFadden (10) and Small (11, 12) respectively. These models are consistent with RUM, but their abilities to identify similarities among choice alternatives are limited.

The limitation of the NL model can be demonstrated with the following example. Suppose we use a NL structure to model the choice from a set of six travel mode alternatives. This choice set contains three automobile modes:

A1 --- Drive Alone,
A2 --- Shared Ride Two Persons in Car, and
A3 --- Shared Ride Three or More Persons in Car,

4-23-4

and three transit modes:

T1 --- Walk Access to Transit,
T2 --- Park-and-Ride Access to Transit, and
T3 --- Kiss-and-Ride Access to Transit.

One natural NL model structure would be to group the three automobile modes into a composite auto alternative and the three transit modes into a composite transit alternative. Under such a circumstance, the NL model structure will allow one parameter to represent the similarity among the three auto modes and another parameter for the three transit modes. In other words, one single correlation is assumed for the three auto modes, and another one for all three transit modes. Differential similarity among modes within each composite alternative is not possible for this structure, despite the fact that, for example, the correlation pattern between walk access and park-and-ride access to transit may be quite different from that between park-and-ride and kiss-and-ride access to transit. Independence between auto modes and transit modes has to be assumed on the marginal choice level. This independence assumption on the marginal level is another conceptual problem in NL. Consider a situation in which the park-and-ride transit mode and the three automobile modes are using the same set of access links in the minimum paths. It would be unrealistic to assume that the park-and-ride transit mode is independent of automobile modes. Thus, a more flexible model is needed.

Ordered Logit (OL) is another model structure that is able to partially relax the IIA assumption in MNL. This structure has been useful for modeling situations where choice decisions are made incrementally. Examples of such decisions are the number of cars to be owned by a household, or the number of children to be raised in a family. With these choice situations, OL is able to capture the similarity among adjacent alternatives. Independence is assumed among the disjoint alternatives.

This paper presents a new structure for probabilistic discrete choice models. This model is derived from assuming choices between paired combinations of alternatives; thus, it is called the paired-combinatorial logit (PCL) model. The PCL model belongs to the family of McFadden's Generalized Extreme Value (GEV) model structures and is consistent with the paradigm of RUM. Unlike NL and OL, PCL is able to identify the similarity between any binary pair of alternatives in the choice set. Thus, it is conceptually more flexible than the NL and OL structures. This model also is considered more practical than MNP and CCL models since this model is mathematically tractable and its choice probabilities and elasticity measures can be computed without numerical integration.

The remainder of this paper is organized as follows. Section 2 presents the analytical formula of the model. Section 3 presents a two-dimensional formulation of the model. Section 4 demonstrates an empirical implementation of the model. Section 5 concludes this paper and suggests future research directions.

2. ANALYTICAL FORMULA

2.1 Deriving the Model

In a significant generalization of MNL, McFadden (10) has derived the GEV model. The GEV model is defined as follows:

Let $G(Y_1, Y_2 \ldots Y_n)$, for $Y_1, Y_2 \ldots Y_n \geq 0$, be a function with the following properties:

1. G is non-negative.

2. G is homogeneous of degree one; that is $G(\alpha Y_1, \alpha Y_2 \ldots \alpha Y_n) = \alpha\, G(Y_1, Y_2, \ldots Y_n)$.

3. $\lim_{Y_k \to \infty} G(Y_1, Y_2 \ldots Y_n) = \infty$, for $k = 1, 2, \ldots n$.

4. The l_{th} partial derivative of G with respect to any combination of l distinct Y_k's, $k = 1, 2, \ldots n$, is non-negative if l is odd and nonpositive if l is even.

If G satisfies these conditions and $G_i\,(Y_1, Y_2, \ldots Y_n)$ denotes $\partial G/\partial Y_i$, $i = 1, 2, \ldots n$, then

$$P(i) = \frac{e^{Vi}\, G_i(e^{V1}, e^{V2}, \ldots e^{Vn})}{G\,(e^{V1}, e^{V2}, \ldots e^{Vn})} \qquad (1)$$

defines the GEV model. In Equation (1), V_i is the systematic utility component of alternative i. V_i often is assumed as a linear function of parameters ß and systematic attributes X_i,

$$V_i = ß'\, X_i \qquad (2)$$

McFadden showed that the choice model defined by Equation (1) is consistent with RUM.

In this paper, I propose to substitute the following G function into McFadden's GEV model

$$G(Y_1\, Y_2 \ldots Y_n) = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (1-\sigma_{ij})\, [Y_i^{1/1-\sigma ij} + Y_j^{1/1-\sigma ij}]^{\,1-\sigma ij}, \text{ where } 0 < \sigma_{ij} < 1. \qquad (3)$$

This G function contains N!/(N-2)!2! terms, which are all the possible paired combinations in the choice set of N alternatives.

It can be shown that this G function satisfies the four conditions laid out by McFadden (15) and

$$G_i = \sum_{j \neq i} (1-\sigma_{ij}) \, [Y_i^{1/(1-\sigma ij)} + Y_j^{1/(1-\sigma ij)}]^{-\sigma ij} \, Y_i^{\sigma ij/(1-\sigma ij)} \tag{4}$$

Substituting Equations (3) and (4) into (1), we can obtain the PCL model as Equation (5).

$$P(i) = \frac{\sum_{j \neq i} e^{Vi/(1-\sigma ij)} (1-\sigma_{ij}) \, [e^{Vi/(1-\sigma ij)} + e^{Vj/(1-\sigma ij)}]^{-\sigma ij}}{\sum_{k=1}^{N-1} \sum_{l=k+1}^{N} (1-\sigma_{kl}) \, [e^{Vk/(1-\sigma kl)} + e^{Vl/(1-\sigma kl)}]^{1-\sigma kl}} \tag{5}$$

The PCL model can be rewritten as

$$P(i) = \sum_{j \neq i} P(i|ij) \, P(ij) \tag{6}$$

where

$$P(i|ij) = \frac{e^{Vi/(1-\sigma ij)}}{e^{Vi/(1-\sigma ij)} + e^{Vj/(1-\sigma ij)}} \tag{7}$$

is the conditional probability of choosing alternative i given the chosen binary pair (ij), and

$$P(ij) = \frac{(1-\sigma_{ij})[e^{Vi/(1-\sigma ij)} + e^{Vj/(1-\sigma ij)}]^{1-\sigma ij}}{\sum_{k=1}^{N-1} \sum_{l=k+1}^{N} (1-\sigma_{kl}) \, [e^{Vk/(1-\sigma kl)} + e^{Vl/(1-\sigma kl)}]^{1-\sigma kl}} \tag{8}$$

is the marginal probability for the binary pair (ij).

2.2 Elasticities

The disaggregate direct elasticity implied by the PCL model can be derived as

$$E_{ii} = \frac{\partial P(i)}{\partial X_i} \frac{X_i}{P_i} = ß \, \{-P(i) + \sum_{j \neq i} P(ij) \, P(i|ij)[1-\sigma_{ij}P(i|ij)]/[P(i)(1-\sigma_{ij})]\} \, X_i \tag{9}$$

and the cross elasticity

$$E_{ij} = \frac{\partial P(i)}{\partial X_j} \frac{X_j}{P_i} = ß \, \{P(j) + \frac{\sigma_{ij}}{1-\sigma_{ij}} [\frac{P(ij)P(i|ij)P(j|ij)}{P(i)}]\} X_j \tag{10}$$

Setting all σ_{kl}'s to zero, the PCL reduces to the standard MNL model and the associated elasticities become

$$E_{ii} = ß[1-P(i)] \tag{11}$$

$$E_{ij} = -ß\, P_j\, X_j \tag{12}$$

2.3 Difference Between NL and PCL Models

The NL model is not a subset of the PCL model. The major difference between these two model structures lies in their respective ability to identify similarity among alternatives. Let's take the six-mode choice situation presented in Section 1 as an example. For the NL formulation, the similarity structure among the six alternatives is defined as in Table 1.

Similarities between auto modes and transit modes are forced to zero because a MNL structure is used on the marginal choice level. The three similarities among auto modes are forced to one parameter σ_A and the three among transit modes are forced to σ_T. Only two similarity parameters are identified by the NL structure. As for PCL structure, all 15 (6!/(4!2!))similarities are identifiable. Now suppose we force these 15 similarities in a PCL to the three values (i.e. 0, σ_A, σ_T) as set by the NL model in Table 1. Then the restricted PCL model will become

$$P(1) = \frac{e^{V1/(1-\sigma A)} [\sum_{j=2}^{3} {}^{(1-\sigma A)} (e^{V1/(1-\sigma A)} + e^{Vj/(1-\sigma A)})^{-\sigma A}] + \sum_{j=4}^{6} e^{V1}}{\sum_{i=1}^{2} \sum_{j=i+1}^{3} {}^{(1-\sigma A)} (e^{Vi/(1-\sigma A)} + e^{Vj/(1-\sigma A)})^{1-\sigma A} + \sum_{i=4}^{5} \sum_{j=i+1}^{6} {}^{(1-\sigma T)} (e^{Vi/(1-\sigma T)} + e^{Vj/(1-\sigma T)})^{1-\sigma T} + \sum_{i=1}^{3} \sum_{j=4}^{6} (e^{Vi} + e^{Vj})} \tag{13}$$

which corresponds to the NL model

$$P(1) = \frac{e^{V1/(1-\sigma A)}}{\sum_{j=1}^{3} e^{Vj/(1-\sigma A)}} \; \frac{[\sum_{j=1}^{3} e^{Vj/(1-\sigma A)}]^{1-\sigma A}}{[\sum_{j=1}^{3} e^{Vj/(1-\sigma A)}]^{1-\sigma A} + [\sum_{j=4}^{6} e^{Vj/(1-\sigma T)}]^{1-\sigma T}} \tag{14}$$

Equation (13) will be identical to (14) only when $\sigma_A = \sigma_T = 0$, which is the case for the MNL structure.

Although NL and PCL are different in structure, both models can solve the red bus/blue bus paradox. Let's look at the NL case first. Suppose alternative 2 (red bus) and 3 (blue bus) are highly correlated (i.e. $\sigma_{23} \sim 1$), and alternative 1 (auto) is independent of alternative 2 and 3 (i.e. $\sigma_{12} = \sigma_{13} = 0$). Then the choice probabilities

$$P(auto) = P(1) = e^{V1}/[e^{V1} + (e^{V2/(1-\sigma 23)} + e^{V3/(1-\sigma 23)})^{1-\sigma 23}]$$

$$P(red\ bus) = P(2) = e^{V2}\ [1 + e^{(V3-V2)/(1-\sigma 23)}]^{-\sigma 23}/[e^{V1} + (e^{V2/(1-\sigma 23)} + e^{V3/(1-\sigma 23)})^{1-\sigma 23}]$$

$$P(\text{blue bus}) = P(3) = e^{V3}\,[1 + e^{(V2-V3)/(1-\sigma 23)}]^{-\sigma 23}/[e^{V1} + (e^{V2/(1-\sigma 23)} + e^{V3/(1-\sigma 23)})^{1-\sigma 23}]$$

As $\sigma_{23} \rightarrow 1$,

$$P(1) = e^{V1}/[e^{V1} + e^{\max(V2,\ V3)}]$$

$$P(2) = \begin{cases} e^{V2}/[e^{V1} + e^{V2}] & \text{if } V_2 < V_3 \\ .5\ e^{V2}/[e^{V1} + e^{V2}] & \text{if } V_2 = V_3 \\ 0 & \text{if } V_2 > V_3 \end{cases}$$

Let $V_1 = V_2 = V_3 = 0$, then $P(1) = 1/2$ and $P(2) = P(3) = 1/4$, which is the desired property of NL model when predicting red bus/blue bus choice probabilities. Let's look at the PCL case. The choice probabilities

$$P(\text{auto}) = P(1) = 2e^{V1}/[2e^{V1} + e^{V2} + e^{V3}]$$

$$P(\text{red bus}) = P(2) = e^{V2}/[2e^{V1} + e^{V2} + e^{V3}]$$

$$P(\text{blue bus}) = P(3) = e^{V3}/[2e^{V1} + e^{V2} + e^{V3}]$$

Let $V_1 = V_2 = V_3 = 0$, then $P(1) = 1/2$ and $P(2) = P(3) = 1/4$, which also satisfies the desired property of relaxing IIA.

2.4 CALIBRATING PCL MODELS

Maximum likelihood estimation, due to its ability to produce efficient and consistent estimators, is the most accepted method for calibrating discrete choice models. MNL and MNP models have been calibrated by the full information maximum likelihood (FIML) method in numerous studies. NL models frequently are calibrated by the limited information maximum likelihood (LIML) method. In the LIML method, the parameters in the NL model are calibrated sequentially. In the first step, the parameter $\beta/(1-\sigma)$ is calibrated based on the conditional choice model, and the log sums of denominators computed. In the second step the similarity parameter $(1-\sigma)$ is calibrated based on these log sums. The attribute parameters β are then recovered by multiplying $\beta/1-\sigma$ from the first step with $(1-\sigma)$ estimated in the second step. LIML method, computationally speaking, is simpler than the FIML method. However, the LIML method does not yield efficient estimators of β and σ. Cosslett (13) and Chu (14) applied the FIML method to calibrate NL models in the transportation mode choice and the residential location choice context. Both studies found that the parameters estimated from the LIML method are remarkably different from those the FIML parameter estimates. Because the parameter values are sensitive to the calibration method being used, and differences in parameter estimates between alternative estimation methods are so large, beyond the range implied by the asymptotic standard errors, McFadden (2) cautioned the users of the LIML method in interpreting the parameter values, and suggested further research on the numerical and statistical properties of these two methods.

The FIML estimation method is applicable for calibrating the attribute parameters ß and the similarity parameters σ in the PCL model. The calibration can be achieved in two steps. In the first step, we force all σ's equal to zero and calibrate the ß parameters in a MNL model. In the second step, using the calibrated coefficients of ß as a starting point, we can calibrate both ß and σ simultaneously. This method may not necessarily lead to a global maximum of the log likelihood function. However, a local maximum, which is superior to MNL model is guaranteed.

3. TWO DIMENSIONAL FORMULATION OF PCL MODEL

Applying the NL model structure to a multidimensional choice situation has been discussed extensively in Ben Akiva and Lerman (1). Suppose there is a two-dimensional choice situation with 3 mode alternatives and 3 destination alternatives. Then, the NL model structure can be written either:

$$P(md) = P(m|d)\ P(d)$$

or

$$P(dm) = P(d|m)\ P(m)$$

In the first formulation, the similarity matrix among the nine joint choice alternatives is only allowed among the 3 mode alternatives. It is shown in Table 2.

Correlations among the destination alternatives are not allowed in this formulation. When the second formulation is used, only the similarities among the three destination alternatives are allowed as shown in Table 3. The similarities among modes can not be incorporated. One simple PCL model structure allowing similarities among modes and destinations simultaneously would have a similarity pattern as shown in Table 4. The multidimensional PCL model can be written explicitly as:

$$P_{A1} = \frac{NUM}{DEN}$$

Where

$$NUM = e^{VA1}\ (1-\sigma) \sum_{m\neq 1} [1 + e^{(VAm-VA1)/(1-\sigma)}]^{-\sigma} +$$

$$e^{VA1}\ (1-\rho) \sum_{d\neq A} [1 + e^{(Vd1-VA1)/(1-\rho)}]^{-\rho} +$$

$$\sum_{d\neq A} \sum_{m\neq 1} e^{VA1}\ [1 + e^{Vdm}\ ]$$

and

$$DEN = (1-\sigma) \sum_{m=1}^{3} (e^{VAm/(1-\sigma)}+e^{VBm/(1-\sigma)})^{1-\sigma}+(e^{VAm/(1-\sigma)}+e^{VCm/(1-\sigma)})^{1-\sigma}+(e^{VBm/(1-\sigma)}+e^{VCm/(1-\sigma)})^{1-\sigma} +$$

$$(1-\rho)\sum_{d=A}^{C}(e^{Vd1/1-\rho}+e^{Vd2/1-\rho})^{1-\rho}+(e^{Vd1/1-\rho}+e^{Vd3/1-\rho})^{1-\rho}+(e^{Vd2/1-\rho}+e^{Vd3/1-\rho})^{1-\rho}+$$

$$4\sum_{m=1}^{3}\sum_{d=A}^{C} e^{Vdm}$$

This PCL model is more complicated than the standard NL model. With the current speed of electronic computers, this model, like MNP, may be too expensive for production purposes. However, this model structure definitely is useful for detecting the co-existence of mode similarities and destination similarities in an effective manner. In the future, when computer speed is enhanced, this model structure will become more practical.

4. EMPIRICAL APPLICATION

The purpose of this section is to demonstrate that a PCL model can be calibrated by the FIML method and that it indeed provides an alternative model structure to MNP and NL when IIA is an important issue in the development of a MNL model. Due to the limited computer resources available, the data set used for this empirical study is very small. It contains 100 individual CBD workers, extracted from the 1970 Chicago census survey. Each individual is associated with four travel modes:

1. Private Automobile,
2. Rapid Transit,
3. Commuter Rail, and
4. Bus.

Using this small data set, a MNL and three PCL models were calibrated by a FORTRAN program. All four models are assumed with a common specification of utility function:

$$V_i = \begin{array}{c}\text{Mode Bias for}\\ \text{Alternative } i\end{array} + \beta_1 \begin{array}{c}\text{One Way Total}\\ \text{Travel Time}\\ \text{in Minutes}\end{array} + \beta_2 \begin{array}{c}\text{Annual Housing}\\ \text{and Transportation}\\ \text{Cost in Dollars}\end{array}$$

The estimation results of the MNL model and the PCL models are reported in Table 5. The equal share model for these 100 cases has the log likelihood -138.6. The market share model for these 100 cases has the log likelihood -136.6. The likelihood ratio test indicates that the market share model does not reject the equal share model at the .05 level (i.e. (138.6 - 136.6) * 2 = 4.0 < 7.82).

When calibrating PCL models, we used the MNL model as the starting point and estimated the most flexible PCL model that calibrates the similarity coefficients for all alternative pairs. Comparing this PCL model with the MNL model we can observe that the magnitude of the β coefficients are extremely close. In fact, all β coefficients in the PCL model are within the 90% confidence intervals implied by the standard errors estimated in the MNL model. This indicates that the MNL model would estimate the β parameters quite accurately even if MNL is an incorrect model structure. The improvement in the log likelihood is very small. The likelihood ratio test indicates that the MNL model cannot be rejected in favor of the PCL model.

Examining the similarity matrix, we found all the similarity coefficients are insignificant. The similarity between the auto mode and of the three public transit modes is negative, a result which is prohibited by our definition of σ. This phenomenon is attributed to the missing attributes effect in the model. More complete specification of the utility function may change these negative similarity coefficients to positive.

The second PCL model was estimated by forcing all the negative similarity coefficients in the previous model to zero, allowing correlations among the three transit modes only. The results are also reported in Table 5. the only significant similarity coefficient, according to the t statistics, is the pair of rapid transit and commuter rail modes. The similarity coefficients between the two rail modes and the bus mode are insignificant. When comparing the similarity coefficients from one model to the other we found that the similarity coefficients changed substantially. The similarity between rapid transit and commuter rail, for instance, has increased from 0.63 to 0.93. Thus, the estimates of similarity coefficients are highly sensitive to the specification of the similarity structure. The third PCL model was estimated by forcing all similarity coefficients, except the one between rapid transit and commuter rail, to zero. The results, as reported in Table 5, shows a new similarity coefficient of 0.35, quite different from those obtained in the other two PCL models.

Based on Table 5, we can conclude that the sample size of 100 cases is too small to detect the statistical difference between the MNL and the PCL models. The β coefficients are remarkably similar in either one of the PCL structures. The estimates of the similarity coefficients are quite unstable. However, the models indicate that there does exist a significant correlation between the random utilities of rapid transit and commuter rail modes. And among the three transit modes, there exists some kind of varied similarities which cannot be captured by the NL model.

Because the estimate of the similarity coefficient is very sensitive to the specification of the similarity structure, it was decided to further explore how this sensitivity will affect the predicted choice probabilities and the implied elasticities. A typical household with the following characteristics:

One Way Travel Time --- 20 minutes (Automobile),
25 minutes (Rapid Transit),
35 minutes (Commuter Rail),
45 minutes (Bus),

Annual Housing Cost --- 2500 dollars,

One Way Travel Cost --- 2.6 dollars (Automobile),
0.6 dollars (Rapid Transit),
1.3 dollars (Commuter Rail),
0.5 dollars (Bus).

was assumed and the choice probability of each mode predicted by these four models for this household is reported in Table 6. From this table we can see that MNL model tends to underestimate the usage of automobiles and overestimate the usage of transit. The predicted choice probabilities are similar between those produced by PCL #2 and PCL #3. However, generally speaking, the more restrictions we have imposed on the similarity structure the higher is the predicted transit share. Between PCL #2 and PCL #3, the attribute parameters, β, are very close, but the similarity coefficients, σ, are very different (0.35 vs. 0.62). Even with this difference in place, the predicted choice probabilities are still very close with each other, which suggests an insensitivity of choice probabilities to the similarity coefficients as long as the major similarities are already accounted for.

The direct elasticities with respect to travel time and travel cost implied by these four models are reported in Table 7. In this table we can see that, for the case of travel time, in each column, automobile mode generally has the smallest elasticities with magnitude being less than one (in absolute value). This suggests that auto travel demands are inelastic to travel time whereas transit demands are elastic to time. In other words, as travel time increases, auto users will not very likely change to transit. But transit riders are more likely to switch to auto.

When comparing the elasticities of automobile mode across alternative model structures, we can see that MNL model is associated with larger values of elasticities (in absolute value) in auto travel demand. That means that MNL model tends to overstate the sensitivity of automobile usage as compared to the PCL model. Suppose PCL is the correct model structure, using MNL model to analyze transportation policy impacts may result in an overestimation of transit usage and underprediction of auto usage.

5. CONCLUSION

This paper presents a new model structure --- paired-combinatorial logit (PCL) --- as an alternative to MNP and NL structures in relaxing the IIA property of MNL models PCL is more flexible than NL in the sense that no hypothesis on the "decision-tree" is needed prior to model calibration and the random utilities can be assumed with any reasonable pattern of correlation. MNL is a special case of PCL in which all similarity coefficients are set to zero. The NL model cannot be obtained as a special case of PCL. PCL is comparable to a special class of MNP in which all random utilities are assumed with identical variance. PCL is more practical than MNP since PCL is mathematically tractable. The probability function and the elasticity measures are all in closed form. Since MNP requires numerical integration, or Clark's approximation (the accuracy of which is a serious concern to model users), PCL appears to be an attractive alternative that warrants consideration. A two-dimensional formulation of PCL is presented in this paper. This two-dimensional structure is much more complicated than the popular NL structure. However, PCL is able to identify the similarity coefficients in both dimensions simultaneously.

Three PCL models were calibrated with 100 cases of mode choice data extracted from 1970 Chicago Census. These models were compared against a MNL model. The PCL models were not able to reject the MNL model based on likelihood ratio statistics. However, the similarity between rapid transit and commuter rail modes appears to be significant. Comparing MNL and the calibrated PCL models, we can conclude that the attribute parameters calibrated for MNL models are close to those of PCL models. In other words, MNL model would estimate the attribute parameters quite accurately even if MNL is the incorrect model structure. However, in terms of predicting choice probabilities or estimating demand elasticities, MNL model tends to overpredict the usage of automobiles and overstating the sensitivity to travel time and travel cost of automobile travel.

Several future research directions can be identified. First, calibrating a PCL model with larger sample sizes (e.g. 500-1000 cases) and a more complete specification would be a worthwhile effort. From this study, the PCL models are not significantly different from MNL, most likely due to the small sample size used in estimation. Second, more detailed comparisons between the PCL model and MNP and NL models can be carried out. With today's computer speeds, the PCL structure may still be too complicated for production purposes; however, PCL definitely would be a useful tool for identifying violations of IIA (or identifying desirable NL tree-structure) in the calibration stage of mode choice models. Third, implementation of multi-dimensional PCL models may be another area of fruitful research.

TABLE 1: SIMILARITY MATRIX FOR NL MODEL STRUCTURE

		1 D A	S R 2	S R 3+	W L K	P N R	K N R
1.	DA	1					
2.	SR2	σ_A	1				
3.	SR3+	σ_A	σ_A	1			
4.	WLK	0	0	0	1		
5.	PNR	0	0	0	σ_T	1	
6.	KNR	0	0	0	σ_T	σ_T	1

TABLE 2: NL MODEL WITH MODE AS CONDITIONAL CHOICE

		A	A	A	B	B	B	C	C	C	--- Destination Alternative
		1	2	3	1	2	3	1	2	3	--- Mode Alternative
A	1	1									
A	2	σ	1								
A	3	σ	σ	1							
B	1	0	0	0	1						
B	2	0	0	0	σ	1					
B	3	0	0	0	σ	σ	1				
C	1	0	0	0	0	0	0	1			
C	2	0	0	0	0	0	0	σ	1		
C	3	0	0	0	0	0	0	σ	σ	1	

TABLE 3: NL MODEL WITH DESTINATION AS CONDITIONAL CHOICE

		A	A	A	B	B	B	C	C	C
		1	2	3	1	2	3	1	2	3
A	1	1								
A	2	0	1							
A	3	0	0	1						
B	1	ρ	0	0	1					
B	2	0	ρ	0	0	1				
B	3	0	0	ρ	0	0	1			
C	1	ρ	0	0	ρ	0	0	1		
C	2	0	ρ	0	0	ρ	0	0	1	
C	3	0	0	ρ	0	0	ρ	0	0	1

TABLE 4: PCL MODEL SIMILARITY STRUCTURE

	A 1	A 2	A 3	B 1	B 2	B 3	C 1	C 2	C 3
A 1	1								
A 2	σ	1							
A 3	σ	σ	1						
B 1	ρ	0	0	1					
B 2	0	ρ	0	σ	1				
B 3	0	0	ρ	σ	σ	1			
C 1	ρ	0	0	ρ	0	0	1		
C 2	0	ρ	0	0	ρ	0	σ	1	
C 3	0	0	ρ	0	0	ρ	σ	σ	1

TABLE 5: COMPARISON BETWEEN PCL AND MNL MODELS

	MNL	PCL #1	PCL #2	PCL #3
Time	-.075 (-4.2)	-.085 (-4.0)	-.064 (-4.0)	-.078 (-4.3)
Log (R+C)	-1.89 (-4.0)	-1.84 (-4.0)	-1.86 (-4.0)	-1.86 (-4.0)
Rapid Transit	-1.31 (-1.2)	-1.41 (-1.0)	-1.32 (-1.0)	-1.37 (-1.0)
Commuter Rail	-.75 (-2.0)	-.75 (-2.0)	-.85 (-2.0)	-.80 (-2.0)
Bus	-.49 (-1.0)	-.49 (-1.0)	-.47 (-1.0)	-.47 (-1.0)
LLO	-138.6	-138.6	-138.6	-138.6
LL*	-130.4	-128.2	-129.3	-130.2

#1 σ_{mn} =

1 (I)			
-.03 (-.04)	1 (I)		
-.06 (-.05)	.63 (.28)	1 (I)	
-.59 (-.6)	.04 (.01)	.09 (.03)	I (I)

#2 σ_{mn} =

1 (I)			
0 (I)	(I)		
0 (I)	.93 (3.5)	1 (I)	
0 (I)	.07 (.2)	.16 (.5)	1 (I)

#3 σ_{mn} =

1 (I)			
0 (I)	1 (I)		
0 (I)	.35 (3.2)	I (I)	
0 (I)	0 (I)	0 (I)	1 (I)

TABLE 6 PREDICTED PROBABILITIES

Direct Elasticity With Respect to Travel Time	PCL #1	PCL #2	PCL #3	MNL
Automobile	0.68	0.64	0.63	0.58
Rapid Transit	0.12	0.15	0.17	0.19
Commuter Rail	0.07	0.09	0.10	0.13
Bus	0.13	0.13	0.10	0.10
Total	1.00	1.00	1.00	1.00

TABLE 7 DIRECT ELASTICITIES

Direct Elasticity With Respect to Travel Time	PCL #1	PCL #2	PCL #3	MNL
Automobile	-0.47	-0.47	-0.58	-0.64
Rapid Transit	-1.99	-1.39	-1.71	-1.51
Commuter Rail	-3.00	-2.14	-2.63	-2.29
Bus	-2.61	-2.63	-3.15	-3.03

Direct Elasticity With Respect to Travel Cost	PCL #1	PCL #2	PCL #3	MNL
Automobile	-0.18	-0.23	-0.24	-0.28
Rapid Transit	-0.19	-0.18	-0.18	-0.17
Commuter Rail	-0.40	-0.38	-0.38	-0.35
Bus	-0.12	-0.16	-0.16	-0.16

REFERENCES

1. Ben Akiva, M. and S. Lerman, (1985), Discrete Choice Analysis: Theory and Application to Travel Demand, Cambridge, MIT Press.
2. McFadden, D., (1981), "Econometric Models of Probabilistic Choice," In Structural Analysis of Discrete Data with Econometric Applications, ed. C.F. Manski and D. McFadden, Cambridge, MIT Press.
3. Schimpeler•Corradino Associates, (1989), "General Planning Consultant Technical Memorandum 89.3.8, Revised Mode-of-Arrival Mode, User Guide," Southern California Rapid Transit District, Los Angeles, California.
4. Barton Aschman Associates, (1984), "Compendium of Technical Memoranda, Transportation Planning and Modeling Services," Southern California Rapid Transit District, Los Angeles, California.

5. Schimpeler•Corradino Associates, (1987), "Regional Mode Choice Model Development Technical Reports," North Central Texas Council of Governments, Dallas.

6. COMSIS Corporation, (1988), "Models of Mode and Occupancy Choice in The Shirley Highway Corridor," Washington, D.C.

7. Daganzo, C., (1979), Multinomial Probit: The Theory and Its Applications to Demand Forecasting, Academic Press, New York.

8. Williams, H.C.L., (1977), "On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefits," Environment Planning A, 9, pp. 285-344.

9. McFadden, D., W.Tye, and K. Train, (1977), "An Application of Diagnostic Tests for the Independence from Irrelevant Alternatives Property of the Multinomial Logit Model," Transportation Research Record, 637, pp. 39-45.

10. McFadden, D., (1978), "Modeling the Choice of Residential Location," In Spacial Interaction Theory and Residential Location, ed. A. Karlgvist et. al. North Holland: Amsterdam, pp. 75-96.

11. Small, K., (1981), "Ordered Logit: A Discrete Choice Model with Proximate Covariance among Alternatives," Working Paper, Department of Economics, Princeton University, Princeton, N.J.

12. Small, K., and C. Hsiao (1985), "Multinomial Logit Specification Tests," International Economic Review, Vol. 26, No. 3, pp. 619-627.

13. Cosslett, S., (1978), Efficient Estimation of Discrete-Choice Models from Choice Based Samples, Ph.D. dissertation, Department of Economics, University of California, Berkeley.

14. Chu, C., (1981), Structural Issues and Sources of Bias in Residential Location and Travel Mode Choice Models, Ph.D. Thesis, Department of Civil Engineering, Northwestern University, Evanston, Illinois.

15. Chu, C., (1989), "Analytical Derivation of Paired Combinatorial Logit Model," Working Paper, Schimpeler•Corradino Associates, Los Angeles.