

# **On Negative Correlation: A Comparison between Multinomial Probit and GEV-based Discrete Choice Models**

Han Dong, Eran Ben-Elia, Cinzia Cirillo, Tomer Toledo, Joseph N. Prashker

Published in *Transportmetrica A: Transport Science* 13(4), pp. 356-379, 2017.

General Extreme Value (GEV) type models like Nested Logit (NL) and Cross Nested Logit (CNL) have gained popularity for their closed-form formulation of the choice probabilities. A key assumption in GEV estimation process is that any correlation between the error terms is necessarily non-negative. No fundamental reason indicates that negative correlations should not occur from a behavioral perspective in the real world. In this paper, we investigate models' outcomes when alternatives exhibit negative correlation. In experiments using synthetic databases, we estimate and validate Multinomial Probit (MNP) models that correctly handle negative correlations and we compare coefficients' estimates and correlations to those obtained with GEV models. A real case study in which choices reveal the presence of negative correlations is also used to assess the performances of the proposed models. Results are obtained with NL, CNL and Mixed Logit models and compared to MNP. The implications for further practices are discussed.

Keywords: GEV Type model; Multinomial Probit Model; Discrete Choice Model; Simulation; Time Use

## **1. Introduction**

Random utility models (RUM) have been developed considerably in the past three decades (Train 2009) and are extensively applied to many travel related behavioral choices. The wide RUM family consists of two main categories: the General Extreme Value (GEV) models based

on the assumption that the errors are type I EV distributed and the Probit model for which the errors are assumed to be multivariate normal. The GEV (McFadden 1978) models include the Multinomial Logit (MNL) and more flexible specifications that allow correlation across choice alternatives while maintaining a closed mathematical form for the choice probabilities. Model formulations that belong to the GEV family include: Nested Logit (Williams 1977); Paired Combinatorial Logit (Chu 1989), Cross-Nested Logit (Vovsha 1997) and General Nested Logit (Wen and Koppelman 2001). The GEV models have been widely applied to model travel mode choice (Hess et al. 2013), spatial location choice (Sener, Pendyala, and Bhat 2011), departure time and route choice (Bekhor, Toledo, and Prashker 2008), and transport networks (Shahhoseini Haghani and Sarvi 2015).

Probit model allows very general error structures ( Ben Akiva and Bolduc 1996; Karac-Mandic and Train 2003; Bhat 2011 ; Daziano and Achtnicht 2013; Daganzo 2014), but the associated choice probabilities requires the computation of multivariate normal distribution functions. In particular, the dimension of the integral depends on the number of correlation terms to be estimated and therefore increases rapidly with number of alternatives. Thus, despite the improvements of estimation techniques (Bhat 2003; Bhat 2001; Daziano and Bolduc 2013; Connors, Hess, and Daly 2014), MNL and other GEV models – mainly Nested and Cross Nested Logit are still those most frequently applied in practical applications involving planning, forecasting and feasibility assessments. However, GEV models are based on a set of specific mathematical properties, one of which is the non-negativity in unobserved correlations. Williams and Ortúzar (1982) presented this condition as rigorous and unambiguous. In reality, there is no fundamental reason why non-positive correlations should not occur also from a behavioral perspective. A negative correlation can appear when an explanatory variable or latent factor is

omitted from the model specification for some reason (if it is not explicitly measured in the data). If this variable has opposite effects on the utilities of two alternatives, their error terms will have a negative correlation. For example, in mode choice, suppose that attitude towards the environment is part of the "true" model, with a positive sign in the utilities of transit, and a negative sign in the car alternatives. But, if this variable is omitted in the specified model, it will generate negative correlations between the transit and car alternatives. Many such examples can be imagined.

In this paper, our motivation is to investigate the bias in the estimation of coefficients and correlation terms deriving from GEV models when errors are negatively correlated. To this scope we design several experiments, we estimate a MNP and we compare the results to those obtained by using NL and CN Logit models. The performance of the three models is assessed with respect to the coefficients' estimates, the ability to recover the correlation among alternatives and the market shares of out-of-sample datasets. The analysis is conducted both on simulated and real data.

The rest of this paper is organized as follows. Section 2 reviews the properties of Nested Logit and Cross Nested Logit under the hypothesis of negative correlation in the error structure. Section 3 presents two experiments designed to assess the degree of the bias in model estimation and prediction, when the assumption of non-negativity is relaxed – the first relates to Nested Logit and the second to Cross-Nested Logit specification. Section 4 presents model estimation results and model validation. All results are compared to those obtained with MNP that has been used to generate the data and that correctly accounts for negative correlation. Section 5 presents a similar comparative analysis for a real case study on leisure activity choices where the performance of Mixed Logit (ML) is also compared to MNP and GEV; the model is also applied

to test the sensitivity to marginal changes in the levels of key attributes. Section 6 presents the conclusions and the implications of our findings on choice based demand models.

## 2. Previous Studies

### 2.1 The GEV theorem and deficiencies in MNL

In GEV theory (McFadden 1978), the probability that a given choice maker ( $n$ ) chooses alternative ( $i$ ) within the choice set ( $C$ ) is:

$$P(i|C) = \frac{y_i G_i(y_1, \dots, y_j)}{\mu G(y_1, \dots, y_j)} = \frac{e^{V_i + \ln G_i}}{\sum_{j \in C} e^{V_j + \ln G_j}} \quad (1)$$

whereby:  $G_i = \partial G / \partial y_i$ ,  $J$  is the number of available alternatives,  $y_i = e^{V_i}$ ,  $V_i$  is the systematic part of the utility function associated with alternative  $i$ , and  $G$  is a non-negative differentiable function which verifies some specific properties.

Any model that can be derived in this way is regarded as a GEV model. This formulation, therefore, defines the family of GEV models. GEV derived models must respect several distinct properties. These properties have no real behavioral intuition but are a mathematical requirement. The properties that the function  $G$  must exhibit are the following:

- (1)  $G \geq 0$  for all positive values of  $y_j \forall j$ .
- (2)  $G$  is homogeneous of degree one i.e. if each  $y_j$  is raised by some proportion  $\rho$ ,  $G$  rises by proportion  $\rho$ . (Ben-Akiva and Francois 1983) showed that this condition can be relaxed to allow any degree of homogeneity.
- (3)  $G \rightarrow \infty$  as  $y_j \rightarrow \infty$  for any  $j$ .

- (4) The mixed partial derivatives of  $G$  exist and are continuous with non-positive even partial derivate. That is,  $G_i \geq 0$  for all  $i$ ,  $G_{ij} = \partial G_i / \partial y_j \leq 0$  for all  $j \neq i$  and non-negative odd mixed partial derivate  $G_{ijk} = \partial G_{ij} / \partial y_k \geq 0$  for any distinct  $i, j$ , and  $k$ , and so on for higher order mixed partials.

These conditions are sufficient so that  $f = e^{-G(e^{y_1}, \dots, e^{y_j})}$  is an absolutely continuous multivariate extreme value distribution function. However, as noted by (Dagsvik 1995), these constraints also imply that the correlations reproduced by a GEV model are necessarily positive.

This general theorem consists of a large family of specifications that includes the MNL itself. MNL's main advantage is in its analytical tractability. However, the hypothesis of errors identically distributed (i.i.d.) causes the property of independence from irrelevant alternatives (IIA), which results in failure to account for similarities between alternatives (Ben Akiva and Lerman 1985). Furthermore, the variance–covariance matrix of the MNL model is homoscedastic.

## ***2.2 Nested Logit and the presence of negative correlations***

Initially proposed by Williams (1977) and Daly and Zachary (1978), the Nested Logit (NL) model is an extension of the MNL model designed to capture correlations among alternatives by partitioning choice sets into different nests. The NL model is designed for choice problems where the alternatives within each nest have correlated error terms; however, error terms between nests remain uncorrelated. Both MNL and NL models are instances of the GEV family. For the MNL model (eq .2):

$$G_i(y_1, \dots, y_j) = \sum_{j \in C} y_j^\mu, \quad P(i | C) = \frac{e^{\mu V_i}}{\sum_{j \in C} e^{\mu V_j}} \quad (2)$$

For the NL model (eq. 3):

$$G_i(y_1, \dots, y_j) = \sum_{m=1}^M \left( \sum_{j \in m} y_j^{\mu_m} \right)^{\frac{\mu}{\mu_m}}, \quad P(i | C) = \frac{e^{\mu_m V_i} \left( \sum_{i \in m} e^{\mu_m V_i} \right)^{\frac{\mu}{\mu_m}}}{\sum_{i \in m} e^{\mu_m V_i} \cdot \sum_{m=1}^M \left( \sum_{i \in m} e^{\mu_m V_i} \right)^{\frac{\mu}{\mu_m}}} \quad (3)$$

where  $\mu$  is the scale factor, and  $\mu_m$  is the parameter associated with nest  $m$ . The ratio of  $\frac{\mu}{\mu_m}$  is the degree of independence or dissimilarity among the alternatives belonging to nest  $m$ . This ratio must be within a particular range for the model to be consistent with utility-maximizing behavior. Following (McFadden 1978), it is possible to assert that the NL model is consistent with utility maximization when  $0 < \frac{\mu}{\mu_m} \leq 1$ . Furthermore, Daganzo and Kusnic (1993) presented

that the correlation between any two alternatives belonging to nest  $m$  is equal to:

$$\text{Corr}(U_{mi}, U_{mj}) = 1 - \left( \frac{\mu}{\mu_m} \right)^2 \quad (4)$$

Börsch-Supan (1990), Kling and Herriges (1995), and Herriges and Kling (1996) provided tests of consistency of NL with utility maximization when the degree of dissimilarity is greater than 1 i.e.  $\frac{\mu}{\mu_m} > 1$ . Train, McFadden, and Ben-Akiva (1987) showed that in this case, consistency with utility-maximizing is appropriate for some specified range of the explanatory variables. Carrasco and Ortuzar (2002) discuss in great details the consistency conditions of

Börsch-Supan and the successive corrections by Kling and Herriges (1995) and Herriges and Kling (1996). They highlights that from a behavioral standpoint a greater degree of substitution between nests than within them makes it impossible to test the hierarchical relationship between the different nesting levels. On the other hand, as noted by Train (2009), a negative value of the degree of dissimilarity i.e.  $\frac{\mu}{\mu_m} < 0$ , is inconsistent with utility maximization and implies that improving the attributes of an alternative (such as lowering its price) can decrease the probability of the alternative being chosen. Finally, when the degree of dissimilarity approaches zero i.e.  $\frac{\mu}{\mu_m} \rightarrow 0$ , NL approaches the “elimination by aspects” model suggested by Tversky (1972). It is interesting to note that the range of  $0 < \frac{\mu}{\mu_m} \leq \sqrt{2}$  is the only proper range that is acceptable in terms of the demands for a correctly specified statistical correlation with corresponding values that lie on the range [-1,1]. Furthermore, when  $1 < \frac{\mu}{\mu_m} \leq \sqrt{2}$ , then there exists a negative correlation between any two alternatives within nest m. This fact certainly cannot be consistent with the GEV theory's assumptions which regard only the possibility of a positive correlation for mathematical reasons.

### ***2.3 Cross Nested Logit and the presence of negative correlations***

The Cross-Nested Logit (CNL) model was also originally proposed by Williams (1977) discussed in terms of its properties using simulated data by Williams and Ortúzar (1982) and further developed by Vovsha (1997). CNL is an extension of the NL model. However, in addition to the choice set being partitioned into nests each alternative may belong to more than one nest. General Nested Logit (GNL) developed by Wen and Koppelman (2001) is a broader

specification than the CNL model. NL is a special case of the GNL model in which the coefficients are binary, either zero or one. Thus an alternative can only belong to one nest. Various formulations for the CNL model have been proposed in the literature (Bierlaire 2006). An adaptation of GNL to model route choice was proposed by Vovsha and Bekhor (1998)

The Paired Combinatorial Logit (PCL) specification is a particular example of the CNL model. PCL is another GEV- type model, proposed by Chu (1989) and later expanded by Koppelman and Wen (2000). It was applied extensively to model route choice by conveniently defining the similarity index (Prashker and Bekhor 1998; Gliebe, Koppelman, and Ziliaskopoulos 1999). In the Nested Logit model all alternatives in a common grouping are similar. In contrast, in the PCL model, each pair of alternatives can have a similarity relationship that is completely independent of the similarity relationship of other pairs of alternatives. This feature is highly desirable for route choice models, since each pair of routes may have different similarities.

Similar to the NL model, CNL also has a GEV generating function and derived probability (eq. 5):

$$G_i(y_1, \dots, y_j) = \sum_{m=1}^M \left( \sum_{j \in C} \left( \alpha_{jm}^{1/\mu} y_j \right)^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

$$P(i | C) = \sum_{m=1}^M \frac{\left( \sum_{i \in m} \alpha_{jm}^{\frac{\mu_m}{\mu}} e^{\mu_m V_j} \right)^{\frac{\mu}{\mu_m}} \cdot \alpha_{im}^{\frac{\mu_m}{\mu}} e^{\mu_m V_i}}{\sum_{n=1}^M \left( \sum_{j \in C} \alpha_{jn}^{\frac{\mu_n}{\mu}} e^{\mu_n V_j} \right)^{\frac{\mu}{\mu_n}} \sum_{j \in C} \alpha_{jn}^{\frac{\mu_n}{\mu}} e^{\mu_n V_j}} \quad (5)$$

where:  $\alpha_{jm} \geq 0, \forall j$ ;  $\sum_{m=1}^M \alpha_{jm} > 0, \forall j$ ;  $\mu_m > 0, \forall m$ ;  $\mu_m \geq \mu, \forall m$ .

However, unlike NL, the correlation between alternatives in overlapping nests is not a simple formula. Papola (2004) proposed a conjecture regarding the approximate structure of this correlation (eq. 6):

$$\text{Corr}(U_{mi}, U_{mj}) = \sum_{m=1}^M \alpha_{im}^{1/2} \alpha_{jm}^{1/2} \left( 1 - \left( \frac{\mu}{\mu_m} \right)^2 \right) \quad (6)$$

Abbe, Bierlaire, and Toledo (2007) presented a proof that this correlation has quite a messy structure which is derived from the joint cumulative distribution function (CDF) of the CNL utilities:

$$\text{Corr}(U_{mi}, U_{mj}) = \frac{6\mu^2}{\pi^2} \iint_{\mathbb{R}^2} x_i x_j \partial_{x_i x_j}^2 F_{\varepsilon_i, \varepsilon_j}(x_i, x_j) dx_i dx_j - \frac{6\gamma^2}{\pi^2} \quad (7)$$

where:  $F_{\varepsilon_i, \varepsilon_j}(x_i, x_j) = \exp \left( - \sum_{m=1}^M \left( \left( \alpha_{im}^{1/\mu} e^{-x_i} \right) + \left( \alpha_{jm}^{1/\mu} e^{-x_j} \right) \right)^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$

This integral has no closed form and must be estimated using numeric procedures. In addition, this correlation (eq. 7) is always positive. Just like in the case of NL, there is no reason to suggest that in reality this assumption should always hold.

### 3. Experiments with synthetic Data

#### 3.1 Rationale

The inherited assumption of non-negative correlations is brought about by mathematical necessities. However, within elaborate nested structures there is no apparent reason why this assumption must hold. Therefore, we decided to put this to the test by creating artificial

correlation structures using synthetic data generation and estimating MNP, and GEV models – NL (Experiment I) and CNL (Experiment II) to measure the obtained bias in the results. MNP unlike GEV can theoretically approximate any correlation structure without bias and should be used whenever the analyst believes that negative correlation could exist. In practice, some restrictions are set on the correlation structure for identification purposes and models' outcomes have some difficulties in their interpretation. For a further discussion of the properties of MNP see (Greene 2008).

### ***3.2 Experiment I***

A sample of 10 files (runs) each with 3,000 synthetic choice observations was created. The sample was created separately for two choice problems: a choice between three alternatives (Experiment Ia) and a choice between four alternatives (Experiment Ib). Each file contained the deterministic utility for each alternative ( $V_i$ ) and the error components ( $\varepsilon_i$ ).

The synthetic utilities – both the deterministic and stochastic parts were computed using a *standard normal distribution*. The alternative specific constant of 1<sup>st</sup> alternative was set to 0 for reasons of normalization. 21 artificial 'true' correlation values ( $\rho$ ) were assumed to vary from -0.95 to 0.95 with 0.095 interval.

For each covariance ( $\rho_k$ ), a variance-covariance matrix was computed. For the three-alternative case the covariance matrix is shown as below:

$$Cov_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho_k \\ 0 & \rho_k & 1 \end{bmatrix}, k = 1, \dots, 21 \quad (8)$$

where:  $Cov_k$  represents the  $k^{th}$  covariance matrix and  $\rho_k$  is true value. In the case of four alternative choice set, the variance-covariance matrix was defined separately for positive and negative correlations:

$$Cov_k = \left\{ \begin{array}{l} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \rho_k & \rho_k \\ 0 & \rho_k & 1 & \rho_k \\ 0 & \rho_k & \rho_k & 1 \end{bmatrix} \text{ if } \rho_k \geq 0, \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\rho_k & \rho_k \\ 0 & -\rho_k & 1 & \rho_k \\ 0 & \rho_k & \rho_k & 1 \end{bmatrix} \text{ if } \rho_k < 0 \end{array} \right\} \quad (9)$$

The vectors of errors of all the alternatives except  $\varepsilon_l$  were multiplied by the Cholesky factorization of each correlation combination in order to transform the matrix into a product of a lower triangular matrix which is important to maintain the stability in the variance-covariance matrix estimation. The chosen alternative was the one with the maximum utility. Thus, for each of the 21 'true' correlations a corresponding vector of choices was matched.

A NL model was estimated with BIOGEME (Bierlaire 2003) for each of the 21 choice vectors in each of the 10 data sets (in total - 210 models). The NL model had a common nest which included all alternatives apart for 1<sup>st</sup> alternative. Figure 1 presents the structure for the three-alternative model and Figure 2 for the four-alternative model:

The NL model was specified according to the following principles:

- (1) The utility functions were specified as:

$$\begin{aligned} U_1 &= \beta_1' V_1 \\ U_i &= \beta_0^i + \beta_i' V_i \end{aligned} \quad (10)$$

where:

$\beta_0^i$  is the alternative specific constant of alternative  $i$  and  $\beta_i'V_i$  is the observed utility components to alternative  $i$ .

(2) The coefficient of the Nest ( $\mu_m$ ) was left to be estimated.

(3) The logit scale ( $\mu$ ) was normalized to 1.

### 3.3 Experiment II

A sample of 10 files (runs) each with 3,000 synthetic choice observations was created using R. The choice was between three alternatives in a similar manner that the data was created in Experiment I.

The artificial correlations were derived from the combinations of the values (0.75, 0.25, -0.25, -0.75) in groups of three. In total  $k=20$  combinations were created. For example, the combination (0.75, 0.75, 0.75) is the first, (0.75, 0.75, 0.25) the second, etc. The covariance matrix was defined as:

$$Cov_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \rho_k^{12} & \rho_k^{13} \\ 0 & \rho_k^{12} & 1 & \rho_k^{23} \\ 0 & \rho_k^{13} & \rho_k^{23} & 1 \end{bmatrix}, \quad k = 1, \dots, 20 \quad (11)$$

whereby:  $\rho_k^{ij}$  is the covariance between alternatives  $i, j$  and  $k$  is the combination's number.

Not all the combinations are viable. In five out of the 20 combinations the Cholesky factorization is invalid. This fact reduced the number of combinations from 20 to 15.

Similar to Experiment I, the vectors of errors of all the alternatives except  $\varepsilon_1$  were multiplied by the Cholesky factorization of each correlation combination. The chosen alternative

was the one with the maximum utility. Thus, for each of the 15 'true' correlations a corresponding vector of choices was matched.

A CNL model was estimated with BIOGEME (Bierlaire 2003) for each of the 15 choice vectors in each of the 10 data sets. The CNL model had a PCL specification of three alternatives, except the first alternative. Each alternative has a shared nest with each of the other two alternatives. Figure 3 depicts the model structure:

The CNL model was specified according to the following principles:

- (1) The utility functions are specified as:

$$\begin{aligned} U_1 &= \beta_1' V_1 \\ U_i &= \beta_0^i + \beta_i' V_i \end{aligned} \tag{12}$$

whereby:  $\beta_0^i$  is the alternative specific constant of alternative  $i$  and  $\beta_i' V_i$  is the overall utility component specified to alternative  $i$ .

- (2) The coefficients of the three nests ( $\mu_m$ ) are left to be estimated.
- (3) The logit scale ( $\mu$ ) is normalized to 1.
- (4) The similarity coefficients ( $\alpha_{im}$ ), were estimated and the sums for each pair are constrained to equal 1.

The estimated correlations of the CNL model are computed using Papola's approximation (eq.6). As Papola's approximation is a conservative estimate of the real correlation, we believe this provides a reasonable estimate of the possible bias compared to the true values.

### 3.4 Normalization of the Covariance Matrix for MNP

In GEV models, the normalization for scale and level occurs automatically with the distributional assumptions that are placed on the error terms. As a result, normalization does not need to be considered for these models. However, with Probit models, normalization for scale and level does not occur automatically. The model should be normalized directly.

The Probit model has  $n$  alternatives, and utility function is expressed as  $U_i = V_i + \varepsilon_i$ ,  $i = 1, \dots, n$ . The vector of errors  $\varepsilon_i$  is normally distributed with zero mean. The procedure proposed by Train (2009) has been applied to normalize the Probit model and assure that all the parameters are identified. The differences with respect to first alternatives are taken, and the error differences is defined as  $\tilde{\varepsilon}_{i1} = \varepsilon_i - \varepsilon_1$ .

The covariance matrix for the vector of error differences take the form

$$\tilde{\Omega} = \begin{pmatrix} \theta_{22} & \cdots & \theta_{2n} \\ \vdots & \ddots & \vdots \\ \theta_{n2} & \cdots & \theta_{nn} \end{pmatrix}$$

where:  $\theta$  is related to the original  $\sigma$ , when the differences are taken against alternative 1. It is showed as follows:

$$\theta_{ab} = \sigma_{ab} + \sigma_{11} - \sigma_{1a} - \sigma_{1b} \quad (13)$$

The matrix is obtained using the  $(n - 1) \times n$  transformation matrix  $M_1$  as

$$\tilde{\Omega} = M_{n-1} \Omega M_{n-1}' \quad (14)$$

where:  $M_{n-1} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & 1 \end{pmatrix}$

## 4. Results

### 4.1 Experiment I

#### 4.1.1 MNP and NL with three alternatives

Figure 4 presents the results of the estimated correlations  $\tilde{\rho}_k$  of the MNP and NL models with three alternatives and the true values. There appears no real difference between the results of the NL model and the true values for positive correlations. However, for negative correlations there is a growing gap between the true value and the estimation as the biased correlation estimates of NL still stay negative. We note that the estimated correlations for the MNP model were basically identical to the true values, as expected.

Table 1 presents the result of the estimated coefficients for the MNP and NL models (averaged over the 10 runs). As the correlation value increases, a smaller scale factor is revealed in the results of NL. When  $\tilde{\rho}_k = 0$ , the scale factor is approximately 1.37, which is consistent with the results in Train (2009). Table 2 shows that MNP does produce consistent results when synthetic data is created with error terms following normal distribution.

#### 4.1.2 MNP and NL with four alternatives

Figure 5 and Figure 6 present the results of the estimation of the correlations of the NL model with four alternatives and the comparison to the true values. Table 2 presents the comparison of the estimated coefficients of the MNP and NL models. The results are averaged over 10 runs.

The correlations shown in Figure 5 and Figure 6, are presented in the form of  $\tilde{\rho}_k$ . Consistent with the results obtained in Experiment Ia, NL model presents bias estimates when negative correlation value exists in the correlation matrix. First, for positive correlation there appears no real difference between the results of the NL model and the true values. While with

negative correlation, there is a growing gap between the true values and the estimation when correlation is decreasing as it can be seen in both figures. The MNP and true values were basically identical.

Table 2 presents the estimated coefficients obtained with the MNP and NL models. Apart for the differences which are attributed to the scale difference between the models, there is no significant difference in the coefficients obtained with NL and MNP. This result is quite remarkable as the correlations clearly show that there is a significant bias in the negative side. However, it seems that the coefficients in the NL model are not influenced by this fact.

## **4.2 Experiment II**

### *4.2.1 MNP and CNL with three alternatives*

Figure 7, Figure 8 and Figure 9 present the comparison between the true covariance and correlation parameters and the PCL model estimates. As noted five out of 20 correlation combinations were not positive semi-definite (i.e. the Cholesky factorization does not exist). These combinations were excluded. Table 3 lists the 15 resulting correlation combinations that were used in the experiment. The estimated correlations in the PCL model were computed according to (Papola, 2004) approximation.

The results show that PCL specification with multiple nests is hard to estimate. Only about 10 out of 100 runs of the model obtained convergence; the log-likelihood function deriving from a PCL specification is highly nonlinear and non-convex, which causes the convergence failures reported. We note that MNP estimates were basically identical to the true values. Figure 7 and Figure 8 show the comparison between converged estimates of CNL model, MNP, and the true values. The results show that the estimated correlation have less bias when all three

correlations are positive ( $k=1, 2, 3, 8$ ). However, when negative correlation value appears, bias can result in the correlation matrix even for correlations with positive values.

Table 4 presents the results for the CNL that converged. Apart for the alternative specific constant, a constant scalar difference can be obtained for most of the results. The results obtained with MNP model are similar to the true value, but are not presented in the paper (can be obtained from the authors by request).

### **4.3 *Model Validation***

In order to calculate the prediction power of the models under analysis, we calculate the market share on out-of-sample datasets. The estimated coefficients based on 2400 observations are applied to the reminder 600 observations. Table 5, Table 6, and Table 7 report the measure of the errors between observed and predicted market shares. In comparing the errors, we conclude that MNP and NL models provide a better fit when compared to CNL. The results indicate that there is not much difference between the predictions of the MNP and NL models for the three-alternative specification. In fact, the NL model has less apparent difference between the true and estimated shares. The results show that both MNP and NL provide reasonable market shares. The difference between true and estimated shares is smaller in NL compared to MNP. In contrast to the results with NL, the validation of CNL shows large differences between the true and estimated market shares. As mentioned in previous section, CNL cannot converge in most of the runs, which leads to the instability in the predictions.

## **5. Evidence from a real case study**

So far our investigation has been based on synthetic data designed specifically for “known” correlation structures. We turn now our attention to a real case study where the primary data

source is extracted from the 2013 American Time Use Survey (ATUS). The ATUS survey has been designed and collected by the Bureau of Labor Statistics on a yearly basis starting from 2003. ATUS questionnaire asks respondents to report their time use together with other information on daily activity episodes including the start and end time of participation, type and location of recorded activity. Socio-demographic information can also be obtained from the survey.

In this study, we consider observations for weekdays from ATUS 2013; 5595 observations are included in the final dataset used for model estimation. Household and individual characteristics, land-use variables and time use information are the main variables extracted from the original dataset. The dependent variable of our discrete choice model is the involvement in leisure activities. Six activity episodes have been selected and categorized according to their locations and types (including computer use for leisure):

- (1) No leisure activities (NO);
- (2) Pure in-home computer use activities (LPC): only choose computer use for leisure activity;
- (3) Pure in-home other leisure activities (LH): only choose in-home leisure activities other than computer use;
- (4) Pure out-of-home leisure activities (LOH): only choose out-of-home leisure activities;
- (5) Multiple in-home leisure and computer use activities (LH&LPC): choose in-home computer use and other in-home leisure activities;
- (6) Multiple in-home and out-home leisure activities (LH&LOH): choose in-home leisure activities without computer use and out-of-home leisure activities.

In addition to the models tested in the synthetic data experiments: MNP, NL, and CNL, there is added value to evaluate the performance of the Mixed Logit Model (ML) (Cardell and Dunbar, 1980; Train, 2009). The ML model is a highly flexible model that can approximate any random utility model (McFadden & Train, 2000) and has been widely applied in research and practice. In this case, the ML model is applied to investigate the negativity of correlations among choices. In the ML model the utility is specified as

$$U_{mj} = \beta' x_{mj} + \kappa_m' z_{mj} + \varepsilon_{mj} \quad (15)$$

where  $x_{mj}$  and  $z_{mj}$  are vectors of observed variables relating to alternative j,  $\beta$  is a vector of fixed coefficients,  $\kappa$  is a vector of random terms with zero mean, and  $\varepsilon_{mj}$  is iid extreme value. The terms in  $z_{mj}$  define the stochastic portion of utility. The unobserved portion of utility is  $\eta_{mj} = \kappa_m' z_{mj} + \varepsilon_{mj}$ , which can be correlated over alternatives depending on the specification of  $z_{mj}$ . The covariance between any two alternatives in nest k is specified as

$$Cov(U_{mi}, U_{mj}) = E(\kappa_k' z_{mi} + \varepsilon_{mi})(\kappa_k' z_{mj} + \varepsilon_{mj}) = \sigma_k \quad (16)$$

A MNP with variance-covariance matrix is also estimated, using in house software coded in R language; as noted several times in this paper MNP is able to correctly recover all types of correlation, including negative correlation if any.

In the NL model, LPC, LH, and LPC&LH are specified in nest B, which contain all home related leisure activities, while LOH and LH&LOH are in nest C, where all the alternatives have an out-of-home leisure episode. It is also conceivable that such correlations also exist between the LH and LH&LOH alternatives, given that they have the common aspect of involving in-home leisure activity. To test for the presence of such correlation, CNL and ML model were

fitted to the data, allowing LH to be shared by two nests. Figure 10 and Figure 11 present the model structures described.

ML, NL, and CNL models were estimated using BIOGEME. The ML correlation matrix attests that there exist negative correlations between alternatives in Nest B and C. The same applies to correlation terms estimated with NL and CNL (shown in Table 8). Unfortunately, these correlation matrices cannot be compared directly with the one obtained by using Probit where notably the correlations are across differences in error terms with respect to the first alternative; for comparison purpose, the covariance matrices are normalized using (eq.14).

Table 9 shows the estimation results obtained by applying the models, along with degree of independence for Nest B and Nest C ( $\mu_B, \mu_C$ ) of both models, factors  $\alpha_{3B}, \alpha_{3C}$  of LH for CNL, and covariance  $\sigma_B, \sigma_C$  for ML. Estimation results are stable and variables maintain their sign and their significance across the three model specifications, with just few exceptions. Surprisingly, the Probit and ML model present worse fit, while NL and CNL produce almost the same value of the final log-likelihood. The nested coefficients  $\mu_B, \mu_C$  are both significant, while the two additional parameters of CNL are not significant.

Consistently with what was conducted for the simulated datasets, we tested the ability of the models in Table 10 to reproduce market share in out-of-samples. We re-estimated the model on about 80% of the observations and we applied the model to the remaining observations. The results show that although ML, MNP, and NL models have a good performance, NL produces better results when compared to MNP. CNL has the most biased results, mainly caused by the failure in reproducing the market share for the alternative LPC&LH.

We finally analyze model elasticity and particularly calculate the effects on LH share caused when increasing of one unit the number of child in the household. Table 11 reports the changes in the aggregate share of LH activity ( $P_{LH}$ ) over the initial value ( $P_{LH}^0$ ):

$$\Delta P_{LH} = \frac{P_{LH} - P_{LH}^0}{P_{LH}^0} \quad (17)$$

where  $P_{LH}$  and  $P_{LH}^0$  are, respectively, the aggregate probabilities of choosing activity LH before and after the variable number of children in the household has been modified. All probabilities are calculated by using sample enumeration (Munizaga et al., 2000).

It appears that ML and MNP models produce similar results, while NL model has different results than the other three. The interpretation is quite straightforward. The NL and CNL models could produce biased modal shifts, when failing to account for correlation across observations and eventually different policy analysis results.

## 6. Conclusions

In this paper we put forward the idea of a possible bias when trying to estimate GEV type choice models in the presence of negative correlations. GEV choice models like Nested Logit and Cross Nested Logit Model have been widely used in the past years. However, modelers hardly ever know in advance the correlation structure of their choice alternatives, and tend to forget the fact that negative error correlation might bias their results. In these cases, MNP or ML, that can overcome the non-negative correlation limitation, should be adopted; however, the simulation assisted estimation is often lengthy and difficult.

To understand the performance of GEV models when negative correlations appear between choices, three experiments are carried out for two of the most common GEV models-

Nested Logit and Cross Nested Logit (Paired Combinatorial). The first two experiments use synthetic data that recreate artificial sets of different correlations in the choice vectors. Based on these datasets we estimate the MNP and GEV models, and we compare their estimates to the true values. An experiment based on the 2013 American Time Use Survey data was considered as a real case study where true values are unknown. However, estimated results obtained using ML model indicate the negative correlations exist between activity choices. The three models were also validated by calculating market shares on out-of-sample observations.

The results with synthetic data (Experiment I and II) reveal that the GEV correlation estimates are biased in the presence of negative correlation, while the MNP estimates of the correlations are practically identical to the true values. In the case of NL, biased estimates of negative correlation have the same patterns for both simple three-alternative case and complex four-alternative case. The results are consistent with the key assumption of GEV model. In the case of CNL, the results from both correlation estimation and validation reveal that the PCL specification fails to estimate the true correlation even under the non-negative conditions. Evidently, more research is required to investigate the CNL model with PCL specification and its failure to achieve convergence.

The results obtained from the real case study attest that MNP, and GEV models produce similar estimates. Negative correlations have been estimated with NL and CNL models; direct comparison with MNP correlations is impossible given that the normalization of probit imposes to work with differences in error terms. While the model fit of NL and CNL is much better than the one obtained with MNP and ML, NL and ML models produce better aggregate choice probabilities when applied to an out-of-sample dataset for validation and when compared with MNP and CNL. Nevertheless, MNP and ML do better than NL in sensitivity analysis when

marginal changes are considered for policy analysis as they properly account for the (negative) correlation across alternatives.

Recently, researchers are working to make it easier to use flexible modeling specifications like MNP by providing more efficient estimation techniques that reduces the computational burden of simulations. This research shows that GEV models, which are notably homoscedastic, could only deal with limited correlation pattern and are not suited for negative correlations. Mixed Logit, which is not limited by the assumptions imposed by GEV is less restrictive. It is suggested that when lacking information on the data structure, more flexible model specifications should be used. However, these models still suffer from a high level of sophistication and expert knowledge is required to verify model identification and correct estimation. Probit and Mixed Logit models have no closed form, estimation is based on simulation and random drawing procedures, and computation time is significantly larger compared to straightforward GEV models. However, improvements made in both hardware and software are reducing this limitation and make flexible models more attractive to practitioners.

The counterintuitive evidence we provided in this paper suggests that more research is needed in understanding the statistical and mathematical properties of discrete choice models. The important lesson for modelers and practitioners is to test many various model specifications with the same dataset including both estimation, and not less important, validation of the model coefficients as well as sensitivity analysis to key parameters.

References:

- Abbe, E., Bierlaire, M., Toledo, T. 2007. "Normalization and correlation of Cross-Nested Logit models." *Transportation Research B*, 41, 795-808.
- Bekhor, S., Toledo, T., & Prashker, J. N. 2008. "Effects of choice set size and route choice models on path-based traffic assignment." *Transportmetrica*, 4(2), 117-133.
- Ben-Akiva, M., Francois, B. 1983. "M $\mu$  homogeneous generalized extreme value model." Working paper, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- Ben Akiva, M., Bierlaire, M. 1999. "Discrete choice methods and their applications to short term travel decisions." *Handbook of Transportation Science*, W. Hall, ed., Kluwer, Dordrecht.
- Ben Akiva, M., Bolduc, D. 1996. "Multinomial probit with a Logit kernel and a general parametric specification of the covariance structure." Working paper, Massachusetts Institute of Technology.
- Bhat, C. R. 2000. "Incorporating observed and unobserved heterogeneity in urban work mode choice modeling." *Transportation Science*, 34, 228-238.
- Bhat, C. R. 2001. "Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model." *Transportation Research B*, 35, 677-693.
- Bhat, C. R. 2003. "Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences." *Transportation Research B*, 37, 837-855.
- Bhat, C. R. 2011. "The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models." *Transportation Research Part B: Methodological*, 45(7), 923-939.
- Bierlaire, M. 2003. "BIOGEME: a free package for the estimation of discrete choice models." In *Swiss Transport Research Conference (No. TRANSP-OR-CONF-2006-048)*. Monte-Verita, Ascona, Switzerland, March 19-21.
- Bierlaire, M. 2006. "A theoretical analysis of the cross-nested logit model." *Annals of operations research*, 144, 287-300.
- Börsch-Supan, A. 1990. "On the compatibility of nested logit models with utility maximization." *Journal of Econometrics*, 43, 371-387.
- Boyd, J., Mellman, J. 1980. "The effect of fuel economy standards on the U.S. automotive market: A hedonic demand analysis." *Transportation Research A* 14, 367-378.

- Brownstone, D., Bunch, D., Train, K. 2000. "Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles." *Transportation Research B*, 34, 315-338.
- Carrasco, J.A. and Ortúzar, J. de D. 2002. "A review and assessment of the nested logit model." *Transport Reviews* 22, 197–218.
- Cardell, S., Dunbar, F. 1980. "Measuring the societal impacts of automobile downsizing." *Transportation Research A*, 14, 423-434.
- Cascetta, E., Nuzzolo, A., Russo, F., & Vitetta, A. 1996. "A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks." In *Proceedings of the 13th International Symposium on Transportation and Traffic Theory* (pp. 697-711). Lyon, France: Pergamon, July 24-26.
- Cascetta, E., Papola, A., Russo, F. R. A. N. C. E. S. C. O., & Vitetta, A. N. T. O. N. I. N. O. 1999. "Implicit availability/perception logit models for route choice in transportation networks." In *World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research* (No. Volume 3).
- Chu, C. 1989. "A paired combinatorial logit model for travel demand analysis." In *Transport Policy, Management & Technology Towards 2001: Selected Proceedings of the Fifth World Conference on Transport Research* (Vol. 4).
- Connors, R. D., Hess, S., & Daly, A. 2014. "Analytic approximations for computing probit choice probabilities." *Transportmetrica A: Transport Science*, 10(2), 119-139.
- Daganzo, C. F., Sheffi, Y. 1977. "On stochastic models of traffic assignment." *Transportation Science*, 11, 253-274.
- Daganzo, C.F. and Kusnic, M., 1993. "Technical Note—Two Properties of the Nested Logit Model." *Transportation Science*, 27(4), pp.395-400.
- Daganzo, C. F. 2014. *Multinomial probit: the theory and its application to demand forecasting*. Elsevier.
- Daly, A.J. and Zachary, S. 1978. "Improved multiple choice models." In D.A. Hensher and M.Q. Dalvi (eds.), *Determinants of Travel Choice*. Saxon House, Westmead.
- Daziano, R. A., & Achtnicht, M. 2013. "Forecasting adoption of ultra-low-emission vehicles using Bayes estimates of a multinomial probit model and the GHK simulator." *Transportation Science*, 48(4), 671-683.

- Daziano, R. A., & Bolduc, D. 2013. "Incorporating pro-environmental preferences towards green automobile technologies through a Bayesian hybrid choice model." *Transportmetrica A: Transport Science*, 9(1), 74-106.
- Dagsvik, J. K. 1995. "How large is the class of generalized extreme value random utility models?" *Journal of Mathematical Psychology*, 39, 90-98.
- Gliebe, J. P., Koppelman, F. S., & Ziliaskopoulos, A. 1999. "Route choice using a paired combinatorial logit model." In 78th meeting of the Transportation Research Board, Washington DC, January 10-14
- Greene, W. H. 2008. *Econometric Analysis 6th Edition*, Pearson/Prentice Hall, Upper Saddle River, N.J.
- Hensher, D., Greene, W. 2001. "The mixed logit model: The state of practice and warnings for the unwary." Working paper, School of Business, The University of Sydney.
- Herriges, J., Kling, C. 1996. "Testing the consistency of nested logit models with utility maximization." *Economic Letters*, 50, 33-39.
- Hess, S., Ryley, T., Davison, L., & Adler, T. 2013. "Improving the quality of demand forecasts through cross nested logit: a stated choice case study of airport, airline and access mode choice." *Transportmetrica A: Transport Science*, 9(4), 358-384.
- Karaca-Mandic, P., & Train, K. 2003. "Standard error correction in two-stage estimation with nested samples." *The Econometrics Journal*, 6(2), 401-407.
- Kling, C., Herriges, J. 1995. "An empirical investigation of the consistency of nested logit models with utility maximization." *American Journal of Agricultural Economics*, 77, 875-884.
- Koppelman, F., Wen, C. 2000. "The paired combination logit model: Properties, estimation and application." *Transportation Research B*, 34, 75-89.
- Luce, R. D., Raiffa, H. 1957. *Games and Decisions*, John Wiley & Sons, New York.
- McFadden, D. 1978. "Modeling the choice of residential location." *Spatial Interaction Theory and Residential Location*, A. Karlqvist, L. Lundquist, F. Snickkars, and J. Weibull, eds., Amsterdam, 75-96.
- McFadden, D., Train, K. 2000. "Mixed MNL models of discrete response." *Journal of Applied Econometrics*, 15, 447-470.

- Munizaga, M. A., Heydecker, B. G., & Ortúzar, J. de D. 2000. "Representation of heteroskedasticity in discrete choice models." *Transportation Research Part B: Methodological*, 34(3), 219-240
- Papola, A. 2004. "Some developments of the Cross- Nested Logit model." *Transportation Research B*, 38, 833-851.
- Prashker, J. N., Bekhor, S. 1998. "Investigation of stochastic network loading procedures." *Transportation Research Record*, 1645, 94-102.
- Revelt, D., Train, K. 2000. "Specific taste parameters and mixed logit." Working paper, Department of Economics, University of California, Berkeley.
- Sener, I. N., Pendyala, R. M., & Bhat, C. R. 2011. "Accommodating spatial correlation across choice alternatives in discrete choice models: an application to modeling residential location choice behavior." *Journal of Transport Geography*, 19(2), 294-303.
- Shahhoseini, Z., Haghani, M., & Sarvi, M. 2015. "Estimation and application of a multi-class multi-criteria mixed paired combinatorial logit model for transport networks analysis." *Transportmetrica B: Transport Dynamics*, 3(1), 59-78.
- Train, K. 1999. "Mixed logit models for recreation demand." *Valuing Recreation and the Environment.*, J. Herriges and C. Kling, eds., Edward Elgar, Northampton, MA.
- Train, K. 2000. "Halton sequences for mixed logit." Working paper, Department of Economics, University of California, Berkeley.
- Train, K. E. 2009. *Discrete choice methods with simulation*. Cambridge university press.
- Train, K., D. McFadden, M. Ben- Akiva. 1987. "The demand for local telephone service: A fully discrete model of residential calling patterns and service choice." *Rand Journal of Economics*, 18, 109-123.
- Tversky, A. 1972. "Elimination by aspects: A theory of choice." *Psychological Review*, 79, 281–299.
- Von-Neumann, J., Morgenstern, O. 1944. *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.
- Vovsha, P. 1997. "The cross-nested logit model: application to mode choice in the Tel- Aviv Metropolitan Area." *Transportation Research Record*, 1607, 6-15.
- Vovsha, P., Bekhor, S. 1998. "The link-nested logit model of route-choice: overcoming the route overlapping problem." *Transportation Research Record*, 1645, 133-148.

- Walker, J. L. 2001. "Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables." Ph. Dthesis, Massachusetts Institute of Technology.
- Walker, J. L., Ben Akiva, M., Bolduc, D. 2004. "Identification of the Logit Kernel (or Mixed Logit) Model." working paper, Massachusetts Institute of Technology.
- Wen, C., Koppelman, F. 2001. "The generalized nested logit model." *Transportation Research B*, 35, 627-641.
- Williams, H.C.W.L. 1977. "On the formation of travel demand models and economic evaluation measures of user benefit." *Environment and Planning 9A*, 285-344.
- Williams, H.C.W.L. and Ortúzar, J. de D. 1982. "Behavioural theories of dispersion and the mis-specification of travel demand models." *Transportation Research 16B*, 167-219.

Table 1: Estimation results - Experiment Ia

$\rho$	Alt1			Alt2			Alt3		
	Asc1	$\beta_{11}$	$\beta_{12}$	Asc2	$\beta_{21}$	$\beta_{22}$	Asc3	$\beta_{31}$	$\beta_{32}$
TRUE	0.000	0.300	1.200	0.800	0.500	1.500	1.500	0.500	1.800
<b>MNP:</b>									
-0.950	0.000	0.295	1.199	0.795	0.481	1.515	1.518	0.506	1.830
-0.855	0.000	0.267	1.222	0.817	0.514	1.470	1.505	0.472	1.809
-0.760	0.000	0.310	1.227	0.796	0.516	1.541	1.526	0.500	1.862
-0.665	0.000	0.287	1.215	0.834	0.504	1.487	1.537	0.480	1.781
-0.570	0.000	0.282	1.222	0.824	0.492	1.532	1.506	0.525	1.828
-0.475	0.000	0.286	1.197	0.785	0.497	1.490	1.486	0.474	1.800
-0.380	0.000	0.321	1.220	0.866	0.486	1.529	1.558	0.485	1.821
-0.285	0.000	0.315	1.175	0.800	0.510	1.525	1.498	0.512	1.786
-0.190	0.000	0.300	1.242	0.818	0.521	1.547	1.542	0.516	1.856
-0.095	0.000	0.315	1.246	0.815	0.492	1.520	1.533	0.526	1.897
0.000	0.000	0.304	1.246	0.838	0.474	1.552	1.553	0.533	1.881
0.095	0.000	0.319	1.213	0.807	0.483	1.539	1.492	0.517	1.832
0.190	0.000	0.305	1.220	0.788	0.481	1.523	1.513	0.534	1.884
0.285	0.000	0.294	1.250	0.811	0.508	1.546	1.536	0.514	1.874
0.380	0.000	0.332	1.224	0.801	0.513	1.554	1.529	0.526	1.869
0.475	0.000	0.292	1.224	0.823	0.500	1.560	1.515	0.536	1.901
0.570	0.000	0.319	1.199	0.814	0.513	1.517	1.518	0.500	1.816
0.665	0.000	0.323	1.213	0.774	0.514	1.536	1.503	0.528	1.861
0.760	0.000	0.294	1.235	0.777	0.497	1.551	1.512	0.515	1.855
0.855	0.000	0.278	1.223	0.816	0.514	1.547	1.529	0.516	1.867
0.950	0.000	0.295	1.248	0.810	0.531	1.577	1.547	0.528	1.901
<b>NL:</b>									
-0.950	0.000	0.436	1.764	1.121	0.705	2.219	2.170	0.732	2.638
-0.855	0.000	0.385	1.774	1.151	0.740	2.109	2.131	0.673	2.572
-0.760	0.000	0.453	1.766	1.108	0.733	2.196	2.141	0.704	2.620
-0.665	0.000	0.416	1.757	1.178	0.721	2.122	2.170	0.681	2.526
-0.570	0.000	0.403	1.716	1.124	0.691	2.141	2.071	0.724	2.517
-0.475	0.000	0.398	1.666	1.062	0.693	2.061	2.027	0.651	2.458
-0.380	0.000	0.435	1.679	1.166	0.665	2.080	2.104	0.653	2.456
-0.285	0.000	0.432	1.607	1.082	0.697	2.071	2.024	0.690	2.409
-0.190	0.000	0.401	1.682	1.086	0.701	2.080	2.052	0.685	2.472
-0.095	0.000	0.422	1.664	1.067	0.647	2.023	2.019	0.687	2.487
0.000	0.000	0.405	1.648	1.085	0.620	2.042	2.027	0.694	2.444
0.095	0.000	0.421	1.606	1.060	0.636	2.024	1.953	0.672	2.400
0.190	0.000	0.396	1.590	1.008	0.620	1.966	1.951	0.680	2.408
0.285	0.000	0.374	1.620	1.043	0.653	1.989	1.974	0.652	2.391
0.380	0.000	0.427	1.564	1.002	0.651	1.977	1.932	0.656	2.346

0.475	0.000	0.373	1.564	1.028	0.638	1.978	1.912	0.674	2.390
0.570	0.000	0.406	1.541	1.038	0.653	1.932	1.933	0.635	2.316
0.665	0.000	0.413	1.558	0.989	0.653	1.953	1.908	0.675	2.366
0.760	0.000	0.374	1.566	0.982	0.623	1.951	1.904	0.648	2.330
0.855	0.000	0.347	1.520	1.003	0.632	1.900	1.887	0.634	2.279
0.950	0.000	0.364	1.547	0.996	0.655	1.938	1.912	0.645	2.326

---

Table 2: Estimation results - Experiment Ib

$\rho$	Alt1			Alt2			Alt3			Alt4		
	Asc1	$\beta_{11}$	$\beta_{12}$	Asc2	$\beta_{21}$	$\beta_{22}$	Asc3	$\beta_{31}$	$\beta_{32}$	Asc4	$\beta_{41}$	$\beta_{42}$
True	0.000	0.300	1.200	0.800	0.500	1.500	1.500	0.500	1.800	0.600	0.600	1.200
<b>MNP:</b>												
-0.950	0.000	0.332	1.197	0.844	0.524	1.527	1.563	0.527	1.844	0.610	0.627	1.199
-0.855	0.000	0.260	1.259	0.841	0.535	1.605	1.585	0.554	1.955	0.615	0.633	1.328
-0.760	0.000	0.308	1.269	0.865	0.531	1.549	1.563	0.508	1.887	0.611	0.626	1.248
-0.665	0.000	0.293	1.233	0.817	0.497	1.505	1.530	0.497	1.825	0.616	0.616	1.167
-0.570	0.000	0.292	1.233	0.831	0.499	1.536	1.527	0.529	1.889	0.543	0.636	1.273
-0.475	0.000	0.299	1.236	0.835	0.509	1.551	1.563	0.485	1.855	0.628	0.618	1.242
-0.380	0.000	0.288	1.202	0.833	0.473	1.443	1.507	0.492	1.769	0.640	0.601	1.152
-0.285	0.000	0.312	1.243	0.851	0.488	1.491	1.552	0.526	1.842	0.618	0.603	1.259
-0.190	0.000	0.320	1.167	0.729	0.520	1.531	1.470	0.525	1.849	0.568	0.594	1.194
-0.095	0.000	0.262	1.224	0.814	0.497	1.533	1.514	0.497	1.828	0.644	0.576	1.200
0.000	0.000	0.320	1.184	0.757	0.524	1.527	1.446	0.505	1.893	0.555	0.608	1.216
0.095	0.000	0.310	1.200	0.759	0.514	1.518	1.472	0.499	1.852	0.543	0.619	1.212
0.190	0.000	0.312	1.210	0.752	0.521	1.566	1.492	0.514	1.845	0.518	0.646	1.262
0.285	0.000	0.315	1.184	0.775	0.516	1.524	1.461	0.498	1.819	0.603	0.581	1.214
0.380	0.000	0.281	1.269	0.878	0.508	1.540	1.579	0.549	1.925	0.621	0.656	1.288
0.475	0.000	0.306	1.235	0.814	0.533	1.564	1.549	0.535	1.869	0.583	0.631	1.270
0.570	0.000	0.306	1.222	0.790	0.515	1.522	1.522	0.504	1.838	0.599	0.629	1.210
0.665	0.000	0.347	1.257	0.831	0.504	1.592	1.556	0.543	1.910	0.585	0.648	1.267
0.760	0.000	0.331	1.273	0.818	0.526	1.591	1.530	0.542	1.915	0.560	0.644	1.304
0.855	0.000	0.308	1.258	0.821	0.540	1.637	1.574	0.521	1.944	0.643	0.628	1.268
0.950	0.000	0.277	1.241	0.778	0.549	1.640	1.534	0.546	1.958	0.563	0.649	1.306
<b>NL:</b>												
-0.950	0.000	0.490	1.760	0.977	0.810	2.344	2.194	0.803	2.741	1.453	0.745	1.431
-0.855	0.000	0.366	1.774	0.931	0.787	2.375	2.130	0.786	2.770	1.383	0.724	1.522
-0.760	0.000	0.437	1.816	1.017	0.780	2.304	2.125	0.741	2.745	1.322	0.755	1.490
-0.665	0.000	0.431	1.789	0.989	0.739	2.262	2.124	0.725	2.641	1.284	0.761	1.445
-0.570	0.000	0.423	1.731	0.988	0.720	2.221	2.059	0.727	2.638	1.122	0.780	1.544
-0.475	0.000	0.417	1.733	1.034	0.713	2.195	2.100	0.671	2.580	1.178	0.764	1.532
-0.380	0.000	0.416	1.720	1.093	0.681	2.054	2.077	0.691	2.473	1.133	0.776	1.484
-0.285	0.000	0.441	1.743	1.086	0.690	2.101	2.098	0.737	2.547	1.052	0.770	1.616
-0.190	0.000	0.452	1.690	1.014	0.741	2.167	2.074	0.732	2.589	0.922	0.815	1.638
-0.095	0.000	0.378	1.762	1.121	0.704	2.165	2.125	0.690	2.534	0.976	0.792	1.654
0.000	0.000	0.460	1.693	1.068	0.726	2.130	2.030	0.691	2.601	0.783	0.853	1.714
0.095	0.000	0.437	1.700	1.044	0.717	2.115	2.038	0.691	2.553	0.771	0.856	1.687
0.190	0.000	0.433	1.661	1.029	0.702	2.111	2.018	0.691	2.477	0.748	0.860	1.694
0.285	0.000	0.439	1.657	1.089	0.698	2.067	2.016	0.670	2.441	0.815	0.804	1.685
0.380	0.000	0.362	1.630	1.085	0.654	1.978	2.004	0.689	2.410	0.823	0.824	1.612

---

0.475	0.000	0.393	1.589	1.031	0.680	1.984	1.962	0.674	2.351	0.774	0.786	1.594
0.570	0.000	0.404	1.602	1.032	0.662	1.963	1.972	0.647	2.366	0.784	0.815	1.566
0.665	0.000	0.434	1.582	1.037	0.622	1.971	1.945	0.661	2.327	0.763	0.794	1.553
0.760	0.000	0.420	1.585	1.001	0.648	1.952	1.894	0.657	2.323	0.711	0.785	1.590
0.855	0.000	0.387	1.560	0.983	0.649	1.960	1.926	0.613	2.283	0.764	0.758	1.524
0.950	0.000	0.352	1.562	0.986	0.667	1.989	1.919	0.657	2.359	0.726	0.786	1.584

---

Table 3: Correlation combinations

k	$\rho_k^{12}$	$\rho_k^{13}$	$\rho_k^{23}$
1	0.75	0.75	0.75
2	0.75	0.75	0.25
3	0.75	0.25	0.25
4	0.75	0.25	-0.25
5	0.75	-0.25	-0.25
6	0.75	-0.25	-0.75
7	0.75	-0.75	-0.75
8	0.25	0.25	0.25
9	0.25	0.25	-0.25
10	0.25	0.25	-0.75
11	0.25	-0.25	-0.25
12	0.25	-0.25	-0.75
13	0.25	-0.75	-0.75
14	-0.25	-0.25	-0.25
15	-0.25	-0.25	-0.75

Table 4: Estimation results of Cross Nested Logit model

<b>k</b>	<b>Alt1</b>			<b>Alt2</b>			<b>Alt3</b>			<b>Alt4</b>		
	<b>Asc1</b>	<b><math>\beta_{11}</math></b>	<b><math>\beta_{21}</math></b>	<b>Asc2</b>	<b><math>\beta_{21}</math></b>	<b><math>\beta_{22}</math></b>	<b>Asc3</b>	<b><math>\beta_{31}</math></b>	<b><math>\beta_{31}</math></b>	<b>Asc4</b>	<b><math>\beta_{41}</math></b>	<b><math>\beta_{41}</math></b>
True	0.000	0.300	1.200	0.800	0.500	1.500	1.500	0.500	1.800	0.600	0.600	1.200
1	0.000	0.417	1.725	0.663	0.874	2.558	1.963	0.779	2.921	0.421	0.996	1.954
2	0.000	0.395	1.576	0.804	0.735	2.231	1.862	0.720	2.556	0.565	0.854	1.731
3	0.000	0.432	1.639	0.979	0.694	2.103	1.981	0.695	2.484	0.728	0.816	1.676
4	0.000	0.390	1.624	1.093	0.652	1.938	2.034	0.615	2.321	0.955	0.806	1.580
5	0.000	0.337	1.641	1.212	0.608	1.815	2.072	0.605	2.201	1.000	0.739	1.481
6	0.000	0.370	1.662	1.309	0.616	1.845	2.306	0.629	2.224	1.387	0.749	1.428
7	0.000	0.467	1.679	1.346	0.629	1.832	2.500	0.632	2.238	1.559	0.728	1.430
8	0.000	0.372	1.661	1.125	0.650	1.991	2.097	0.644	2.397	0.823	0.800	1.586
9	0.000	0.457	1.617	1.187	0.649	1.878	2.077	0.631	2.220	0.956	0.766	1.480
10	0.000	0.441	1.745	1.226	0.669	2.006	2.148	0.681	2.372	0.955	0.877	1.649
11	0.000	0.449	1.616	1.216	0.594	1.838	2.094	0.588	2.224	0.973	0.740	1.475
12	0.000	0.406	1.747	1.198	0.669	2.026	2.143	0.632	2.387	0.958	0.812	1.649
13	0.000	0.495	1.828	1.250	0.722	2.129	2.271	0.699	2.491	0.992	0.904	1.731
14	0.000	0.420	1.740	1.362	0.639	1.901	2.225	0.616	2.200	1.070	0.763	1.570
15	0.000	0.428	1.690	1.334	0.676	1.940	2.239	0.631	2.257	1.041	0.813	1.564

Table 5: Difference in market share predictions - Experiment Ia

$\rho$	MNP			NL		
	Alt1	Alt2	Alt3	Alt1	Alt2	Alt3
-0.95	2.9%	-1.4%	-1.4%	-0.6%	0.4%	0.2%
-0.855	3.4%	-1.6%	-1.8%	0.2%	0.0%	-0.2%
-0.76	2.8%	-1.5%	-1.3%	-0.4%	0.1%	0.3%
-0.665	3.3%	-2.0%	-1.3%	0.4%	-0.7%	0.4%
-0.57	3.2%	-1.4%	-1.8%	0.6%	-0.2%	-0.3%
-0.475	2.6%	-1.0%	-1.7%	0.2%	-0.1%	-0.1%
-0.38	2.6%	-0.9%	-1.7%	0.4%	-0.2%	-0.2%
-0.285	1.9%	0.1%	-2.1%	-0.3%	0.8%	-0.5%
-0.19	2.1%	-0.6%	-1.5%	0.3%	-0.2%	-0.1%
-0.095	1.3%	-0.2%	-1.1%	-0.3%	0.0%	0.3%
0	1.8%	0.0%	-1.8%	0.4%	0.0%	-0.4%
0.095	1.6%	0.3%	-1.9%	0.3%	0.1%	-0.4%
0.19	1.2%	-0.4%	-0.8%	0.2%	-0.9%	0.7%
0.285	0.8%	0.8%	-1.7%	0.0%	0.3%	-0.2%
0.38	0.4%	0.9%	-1.4%	-0.1%	0.2%	-0.1%
0.475	0.3%	1.2%	-1.5%	-0.1%	0.4%	-0.3%
0.57	0.0%	1.6%	-1.5%	-0.3%	0.3%	0.0%
0.665	0.1%	1.0%	-1.1%	0.1%	-0.4%	0.4%
0.76	-0.4%	2.1%	-1.7%	-0.2%	0.7%	-0.6%
0.855	-1.0%	2.5%	-1.5%	-0.3%	0.3%	0.0%
0.95	-0.7%	1.6%	-0.9%	0.0%	-0.4%	0.4%

Table 6: Difference in market share predictions - Experiment Ib

$\rho$	MNP				NL			
	Alt1	Alt2	Alt3	Alt4	Alt1	Alt2	Alt3	Alt4
-0.950	2.2%	2.8%	-0.1%	-4.8%	-0.1%	0.2%	0.3%	-0.5%
-0.855	2.0%	1.7%	-0.4%	-3.4%	-0.1%	-0.4%	-0.1%	0.7%
-0.760	2.3%	1.7%	-1.0%	-2.9%	0.3%	-0.4%	-0.4%	0.4%
-0.665	1.8%	1.8%	-1.4%	-2.2%	-0.2%	0.1%	-0.6%	0.7%
-0.570	1.7%	1.6%	-1.5%	-1.8%	-0.1%	0.1%	-0.6%	0.6%
-0.475	1.2%	1.2%	-0.2%	-2.2%	-0.5%	0.1%	0.6%	-0.2%
-0.380	1.3%	1.5%	-2.0%	-0.9%	-0.3%	0.4%	-0.6%	0.5%
-0.285	1.6%	1.2%	-1.1%	-1.7%	0.1%	0.3%	0.0%	-0.5%
-0.190	1.4%	0.2%	-1.2%	-0.4%	-0.5%	0.1%	0.3%	0.0%
-0.095	2.1%	-0.4%	-1.7%	0.0%	0.5%	-0.5%	-0.1%	0.1%
0.000	2.2%	-0.4%	-2.6%	0.8%	0.3%	-0.1%	-0.8%	0.5%
0.095	1.2%	0.0%	-1.3%	0.2%	-0.3%	0.0%	0.6%	-0.2%
0.190	1.2%	0.4%	-1.8%	0.3%	0.1%	0.3%	-0.1%	-0.3%
0.285	0.8%	0.8%	-2.6%	1.0%	0.0%	0.7%	-0.3%	-0.4%
0.380	-0.1%	1.4%	-2.5%	1.2%	-0.1%	0.6%	-0.5%	0.0%
0.475	-0.4%	0.7%	-1.7%	1.3%	-0.2%	0.0%	0.4%	-0.2%
0.570	-0.3%	1.0%	-2.7%	1.9%	0.1%	0.1%	-0.2%	0.1%
0.665	-1.3%	0.4%	-2.3%	3.3%	-0.5%	-0.4%	0.1%	0.9%
0.760	-0.8%	0.6%	-2.4%	2.6%	0.4%	-0.3%	-0.3%	0.3%
0.855	-1.8%	0.9%	-2.4%	3.3%	-0.1%	-0.2%	0.1%	0.2%
0.950	-1.5%	1.2%	-2.9%	3.2%	0.2%	0.0%	-0.3%	0.1%

Table 7: Difference in market share predictions - Experiment II

k	MNP				CNL			
	Alt1	Alt2	Alt3	Alt4	Alt1	Alt2	Alt3	Alt4
1	0.5%	0.3%	-3.0%	2.2%	-8.8%	7.5%	-10.5%	11.7%
2	-1.5%	1.9%	-3.1%	2.7%	-7.7%	5.2%	-4.7%	7.3%
3	-0.1%	3.0%	-1.4%	-1.6%	-5.8%	6.5%	-9.7%	9.1%
4	0.8%	1.9%	-1.8%	-0.9%	-9.8%	17.5%	-26.0%	18.3%
5	0.6%	2.4%	-0.3%	-2.7%	-4.9%	6.7%	-1.8%	0.1%
6	2.7%	2.8%	-2.7%	-2.8%	-6.0%	12.9%	-8.6%	1.7%
7	2.0%	1.6%	-2.0%	-1.6%	-4.1%	9.1%	-0.1%	-4.9%
8	1.6%	-0.7%	-1.1%	0.2%	-8.1%	10.0%	-6.4%	4.5%
9	1.7%	2.6%	-3.4%	-0.9%	-6.5%	6.1%	-9.6%	9.9%
10	1.9%	1.2%	-2.3%	-0.8%	-4.7%	11.7%	-7.8%	0.9%
11	0.5%	1.8%	-1.7%	-0.6%	-4.9%	5.0%	-6.0%	5.9%
12	0.9%	2.4%	-1.1%	-2.2%	-3.6%	9.9%	-5.3%	-1.0%
13	2.3%	0.1%	-0.9%	-1.6%	-2.9%	10.6%	3.9%	-11.6%
14	2.7%	-0.9%	-0.4%	-1.3%	-6.1%	8.9%	-6.1%	3.3%
15	2.4%	2.3%	-2.8%	-1.9%	-6.2%	16.4%	-11.7%	1.5%

Table 8: Covariance of difference Matrix in real case

$$\tilde{\Omega}_{MNP} = \begin{pmatrix} 2.00 & 0.84 & 1.14 & 1.16 & 1.02 \\ 0.84 & 2.78 & 1.26 & 1.89 & 1.86 \\ 1.14 & 1.26 & 3.95 & 2.76 & 2.83 \\ 1.16 & 1.89 & 2.76 & 5.21 & 4.36 \\ 1.02 & 1.86 & 2.83 & 4.36 & 6.12 \end{pmatrix} \quad \tilde{\Omega}_{NL} = \begin{pmatrix} 3.29 & -1.72 & 1.64 & -1.72 & 1.64 \\ -1.72 & 3.29 & 1.64 & -1.72 & 1.64 \\ 1.64 & 1.64 & 3.29 & 1.64 & 0.76 \\ -1.72 & -1.72 & 1.64 & 3.29 & 1.64 \\ 1.64 & 1.64 & 0.76 & 1.64 & 3.29 \end{pmatrix}$$

$$\tilde{\Omega}_{CNL} = \begin{pmatrix} 3.29 & -2.15 & 1.64 & -8.40 & 1.64 \\ -2.15 & 3.29 & 1.39 & -2.15 & 1.39 \\ 1.64 & 1.39 & 3.29 & 1.64 & 1.37 \\ -8.40 & -2.15 & 1.64 & 3.29 & 1.64 \\ 1.64 & 1.39 & 1.37 & 1.64 & 3.29 \end{pmatrix} \quad \tilde{\Omega}_{MXL} = \begin{pmatrix} 3.01 & 1.36 & 1.64 & 1.36 & 1.64 \\ 1.36 & 3.23 & 1.87 & 1.36 & 1.87 \\ 1.64 & 1.87 & 3.51 & 1.64 & 1.87 \\ 1.36 & 1.36 & 1.64 & 3.01 & 1.64 \\ 1.64 & 1.87 & 1.87 & 1.64 & 3.51 \end{pmatrix}$$

Correlation matrix of NL, CNL, ML

$$Cor_{NL} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2.05 & 0 & -2.05 & 0 \\ 0 & -2.05 & 1 & 0 & -2.05 & 0 \\ 0 & 0 & 0 & 1 & 0 & -0.54 \\ 0 & -2.05 & -2.05 & 0 & 1 & 0 \\ 0 & 0 & 0 & -0.54 & 0 & 1 \end{pmatrix}$$

$$Cor_{CNL} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2.31 & 0 & -6.11 & 0 \\ 0 & -2.31 & 1 & -0.15 & -2.31 & -0.15 \\ 0 & 0 & -0.15 & 1 & 0 & -0.16 \\ 0 & -6.11 & -2.31 & 0 & 1 & 0 \\ 0 & 0 & -0.15 & -0.16 & 0 & 1 \end{pmatrix}$$

$$Cor_{MXL} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -0.19 & 0 & -0.21 & 0 \\ 0 & -0.19 & 1 & 0.13 & -0.19 & 0.13 \\ 0 & 0 & 0.13 & 1 & 0 & 0.12 \\ 0 & -0.21 & -0.19 & 0 & 1 & 0 \\ 0 & 0 & 0.13 & 0.12 & 0 & 1 \end{pmatrix}$$

Table 9: Estimation results in real case (standard errors in parentheses)

	Name	Coefficient			
		MNP	NL	CNL	ML
<b>LPC</b>	Constant	-1.091*	-4.810*	-7.030	-1.800*
		(0.189)	(2.22)	(4.51)	(0.24)
	Higher than BA Degree (Dummy)	0.714	1.010	1.300	0.538*
		(0.3)	(0.65)	(0.94)	(0.29)
	Full-time working status (Dummy)	-0.109	-0.254	0.252	-0.424*
		(0.245)	(0.49)	(1.03)	(0.25)
<b>LH</b>	Constant	2.520*	2.980*	2.960*	3.040*
		(0.092)	(0.12)	(0.14)	(0.31)
	No. of children	-0.030*	-0.142*	-0.140*	-0.147*
		(0.037)	(0.05)	(0.05)	(0.04)
	Full-time working status (Dummy)	-0.758*	-0.674*	-0.702*	-0.744*
		(0.073)	(0.09)	(0.11)	(0.11)
	Age. Senior (Dummy)	-0.362*	-0.673*	-0.708*	-0.476*
		(0.089)	(0.13)	(0.17)	(0.09)
	Age. Adult (Dummy)	-0.372*	-0.863*	-0.888*	-0.817*
		(0.118)	(0.16)	(0.19)	(0.13)
Age. Young (Dummy)	-0.646*	-1.090*	-1.170*	-0.847*	
	(0.153)	(0.22)	(0.28)	(0.16)	
Age. Teen (Dummy)	-0.579*	-1.040*	-1.130*	-0.904*	
	(0.162)	(0.24)	(0.30)	(0.17)	
Higher than BA Degree (Dummy)	-0.156*	-0.434*	-0.422*	-0.335*	
	(0.118)	(0.15)	(0.15)	(0.12)	
<b>LOH</b>	Constant	-1.084*	-1.320*	-1.210*	-0.955*
		(0.153)	(0.61)	(0.56)	(0.19)
	No. of children	-0.209*	-0.320*	-0.313*	-0.278*
		(0.162)	(0.12)	(0.11)	(0.09)
	Single (Dummy)	0.197	0.273	0.295	0.317*
		(0.186)	(0.25)	(0.23)	(0.19)
Age. Teen (Dummy)	0.338*	1.210*	1.150*	0.957*	
	(0.309)	(0.43)	(0.39)	(0.31)	
Age. Young (Dummy)	0.632*	0.768*	0.737*	0.870*	
	(0.268)	(0.38)	(0.34)	(0.26)	
<b>LPC&amp;LH</b>	Constant	0.420*	-0.641	-0.857	0.837*
		(0.1)	(1.03)	(1.5)	(0.19)
	No. of children	-0.098*	-0.287*	-0.310*	-0.247*
		(0.056)	(0.10)	(0.11)	(0.06)
	Race. Black (Dummy)	-0.775*	-1.250*	-1.480*	-0.667*
		(0.161)	(0.49)	(0.60)	(0.16)
	Full-time working status (Dummy)	-1.341*	-1.510*	-1.660*	-1.29*
	(0.115)	(0.25)	(0.34)	(0.14)	
Higher than BA Degree (Dummy)	-0.004	0.605*	0.615*	0.307*	
	(0.167)	(0.35)	(0.37)	(0.17)	
BA Degree (Dummy)	0.557*	0.820*	0.894*	0.548*	
	(0.114)	(0.28)	(0.30)	(0.12)	
Age. Adult (Dummy)	-0.190*	-0.545*	-0.518*	-0.370*	

<b>LOH&amp;LH</b>		(0.139)	(0.25)	(0.27)	(0.14)
	Constant	0.378*	0.603*	0.531*	0.704*
		(0.103)	(0.14)	(0.29)	(0.17)
	No. of children	-0.106*	-0.161*	-0.168*	-0.169*
		(0.054)	(0.07)	(0.07)	(0.05)
	Gender. Female (Dummy)	-0.475*	-0.371*	-0.390*	-0.278*
		(0.086)	(0.10)	(0.13)	(0.09)
	Single	0.359*	0.357*	0.370*	0.313*
		(0.105)	(0.13)	(0.13)	(0.11)
	Higher than BA Degree (Dummy)	-0.581*	-0.857*	-0.808*	-0.819*
		(0.176)	(0.24)	(0.22)	(0.18)
BA Degree (Dummy)	0.025*	-0.242*	-0.217	-0.275*	
	(0.111)	(0.14)	(0.14)	(0.11)	
Age. Adult (Dummy)	-0.603*	-0.427*	-0.414*	-0.495*	
	(0.125)	(0.16)	(0.15)	(0.13)	
	$\mu_B$	0.573*	0.375*		
	$\mu_C$	0.807*	0.927		
	$\alpha_{3B}$		0.143*		
	$\alpha_{3C}$		0.857*		
	$\sigma_B$			-0.281	
	$\sigma_C$			0.233	
	Final Log-likelihood	-6420.89	-4975.36	-4975.02	-6213.43

\*significant, p-value < 0.1

Table 10: Validation results in real case study (predicted market shares)

Alternative	Observed	Predicted value				Difference			
		MNP	NL	CNL	ML	MNP	NL	CNL	ML
NO	9.20%	10.81%	10.86%	12.49%	9.48%	1.61%	1.66%	3.28%	0.27%
LPC	0.89%	3.90%	0.00%	0.00%	1.35%	3.00%	-0.89%	-0.89%	0.45%
LH	64.70%	63.47%	62.78%	70.51%	65.32%	-1.23%	-1.92%	5.81%	0.62%
LOH	3.84%	3.46%	2.78%	5.51%	3.42%	-0.38%	-1.06%	1.67%	-0.42%
LPC&LH	9.38%	8.07%	12.10%	0.41%	9.12%	-1.31%	2.72%	-8.98%	-0.26%
LOH&LH	11.97%	10.29%	11.48%	11.09%	11.31%	-1.68%	-0.50%	-0.89%	-0.67%

Table 11 Policy analysis

Alternative	MNP	NL	CNL	ML
NO	3.49%	18.41%	11.22%	7.67%
LPC	3.70%	37.86%	22.59%	8.80%
LH	0.72%	4.67%	-0.04%	0.01%
LOH	-10.84%	-12.47%	-17.50%	-6.74%
LPC&LH	-2.17%	-38.04%	-8.67%	-5.03%
LOH&LH	-2.37%	0.89%	-4.61%	-1.10%

Figure 1: NL model with three alternatives

Figure 2: NL model with four alternatives

Figure 3: Cross Nested Logit model with four alternatives

Figure 4: Comparison of correlation  $\tilde{\rho}_k$  between NL and MNP model

Figure 5: Comparison of  $\tilde{\rho}_k^{12}$  between NL and MNP model

Figure 6: Comparison of  $\tilde{\rho}_k^{13}$  between NL and MNP model

Figure 7: Comparison of  $\tilde{\rho}_k^{12}$  between CNL and MNP model

Figure 8: Comparison of  $\tilde{\rho}_k^{13}$  between CNL and MNP model

Figure 9: Comparison of  $\tilde{\rho}_k^{23}$  between CNL and MNP model

Figure 10: NL model structure in real case study

Figure 11: CNL and ML model structure in real case study