# MLMNP AND MLMNPB:
# Fortran Programs for Maximum Likelihood Estimation of Linear-in-Parameter Multinomial Probit Models

*by*

**David S. Bunch***

**UCD-ITS-RR-91-14**

**September 1991**

**Institute of Transportation Studies**
**University of California, Davis**
**Davis, California 95616**

***Graduate School of Management**
**308 Voorhies**
**University of California, Davis**
**Davis, California 95616-8609**

**(916) 752-2248**
**Fax: (916) 752-2924**
**E-mail: dsbunch@ucdavis.edu**

# Table of Contents

## 1. Introduction

The Fortran driver programs MLMNP and MLMNPB perform maximum likelihood (ML) estimation for the linear-in-parameters multinomial probit (MNP) model, e.g., models discussed in Bunch (1991) and Bunch and Kitamura (1990), and defined in section 3 below. The programs provide an easy-to-use but flexible input format, and they handle storage allocation to facilitate use of the optimization subroutines described in Bunch, Gay, and Welsch (1991). They follow a standardized framework for ML estimation of a general probabilistic choice model (except for the additional code specific to the MNP model) and could thus be used as prototypyes for other types of discrete choice models, where the MNP parts of the code are replaced with code for the appropriate model.

MLMNP features a variety of capabilities for calculating the regression diagnostics described in Bunch, Gay, and Welsch (1991). The "B" in MLMNPB indicates that the program enforces simple bounds on the parameter values, which is useful for preliminary work with models and data sets that might contain singularities or specification errors (see discussion below). Because of the difficulty of calculating variance-covariance matrix estimates for boundary solutions, MLMNPB does not provide the regression diagnostics and asymptotic t-statistics that are available in MLMNP. Thus, MLMNP assumes that the sequence of iterates remains in the region of the parameter space where the model is well defined.

The method used by both programs to compute the probit integrals is the fixed-order Mendell-Elston (FOME) approximation discussed in Bunch and Kitamura (1990) and Kamakura (1990); the current implementation allows choice sets to contain up to 20 alternatives. Although the FOME approximation has been demonstrated to be superior to Clark's approximation for choice sets of size three or four, there is little evidence regarding the method's behavior for larger choice sets: users estimating models with large numbers of alternatives do so at their own risk.

This report provides background information, instructions for program usage, and examples. (However, the programs themselves are fully commented and could be used without this document.) Section 2 provides a review of linear-in-parameters MNP models, including models with random taste variation, and/or general covariance error structure. Section 3 describes the general framework for estimation of probabilistic choice models. Section 4 gives a description of the input format, and section 5 contains examples. The code and the examples have been submitted for inclusion in the Collected Algorithms of the ACM, along with the submission of Bunch, Gay, and Welsch (1991) to the *ACM Transactions on Mathematical Software*.

## 2. Linear-in-Parameters Multinomial Probit Models

Consider a sample of N decision makers (or, DM's), indexed by n = 1, ..., N, and a universe of J choice alternatives, denoted by the set C, where the choice alternatives in C are identified by nominal labels. For example, in a problem involving choice of transport mode for the commute to work C = {drive alone, carpool, train, bus, ...}. Not every choice may be available to every DM; each DM chooses one alternative from a choice set $C_n$, where $C_n$ is a subset of C containing $J_n$ alternatives. In most models, explanatory data are collected such that each of the $J_n$ alternatives is characterized by NATTR (generic) attributes, which are stored in the NATTR x $J_n$ matrix $X_n$. We assume that choice is governed by the following linear-in-parameters random utility framework:

$$(1) \qquad U_n = X_n^T(\theta + \delta_n) + \mu_n + \varepsilon_n, \quad n = 1, ..., N,$$

where $U_n, \mu_n, \varepsilon_n \in \mathbf{R}^{J_n}$, and $\theta, \delta_n \in \mathbf{R}^{NATTR}$. The vector $U_n$ contains DM n's (unobserved) true utilities for the $J_n$ alternatives, and the observed choice will be the one with the largest utility. The NATTR-vector $(\theta + \delta_n)$ contains the attribute taste weights for DM n, where $\theta$ is the mean taste weight for the population and $\delta_n$ is the (unobserved) random deviation from the mean for household n, where $E(\delta_n) = 0$.

The sum $(\mu_n + \varepsilon_n)$ represents the effect of (unobserved) random errors on utility, where $\mu_n$ is the vector of alternative-specific means for the nominal choices in $C_n$, and $\varepsilon_n$ is the vector of (alternative-specific) random errors with $E(\varepsilon_n) = \mathbf{0}$. These terms may be regarded as including the effect of unobserved attributes of the individual and/or the alternatives, as well as other sources of error. Note than $\mu_n$ is a collection of components from the J-vector $\mu$, the alternative-specific mean vector for all the nominal choices in C.

To get a multinomial probit model, one adds the theoretically appealing assumption that both random terms have multivariate normal distributions:

$$(2) \qquad \delta \sim MVN(0, \Sigma_\delta), \Sigma_\delta \in \mathbf{R}^{NATTR \times NATTR} \text{ and } \varepsilon \sim MVN(0, \Sigma_\varepsilon), \Sigma_\varepsilon \in \mathbf{R}^{J \times J}.$$

The covariance matrix for $\varepsilon_n$ is constructed by extracting the rows and columns of $\Sigma_\varepsilon$ corresponding to the nominal alternatives in $C_n$.

The term $\mu_n$ in equation (1) could also have been represented via an alternative framework which allows dummy variables in $X_n$; the alternative-specific random errors would then be represented by the $\delta$ terms corresponding to the dummies, and no $\varepsilon$ error terms would be required. By using (1) and (2) instead, we choose a slightly different framework that assumes the $\delta$ and $\varepsilon$ terms are always independent; this reduces the number of covariance parameters required and simplifies the analysis of model specifications.

Various model specifications may be formulated by placing restrictions on the random taste covariance matrix $\Sigma_\delta$ and the error covariance matrix $\Sigma_\varepsilon$ in (2). Three types of taste variation are considered: (i) fixed (non-random) tastes with $\Sigma_\delta \equiv 0$, (ii) uncorrelated random tastes with $\Sigma_\delta \equiv D$ (a diagonal matrix), and (iii) correlated random tastes with a general covariance matrix $\Sigma_\delta$.

Now consider the full covariance matrix $\Sigma_\varepsilon$ for error terms involving the alternative-specific nominal variables in C. Two types of $\varepsilon$'s are considered: (i) IID errors with $\Sigma_\varepsilon = kI$, where k is a scale constant, and (ii) non-IID errors with a "general" $\Sigma_\varepsilon$. In the case where alternatives are defined exclusively by their generic attributes and no nominal labels are used, it is appropriate to use IID random errors. Scaling in the probit model is arbitrary, and we will usually assume that k=1 for the IID error specifications. However, the specification of a "general" $\Sigma_\varepsilon$ requires special consideration, since all J(J+1)/2 parameters are not estimable.

Briefly, the estimability problem arises because observed choices may always be identified based on *differences* of utilities rather than on the utilities themselves. This means that covariance parameters associated with differences in error terms can be estimated, but that the covariance parameters for the error terms themselves can not. Specifically, the order of a covariance matrix for error differences is (J-1) rather than J, i.e., one less than the original covariance matrix. This means that there are only J(J-1)/2 parameters that can be estimated rather than the original J(J+1)/2. In addition, one parameter is used to fix the scaling of the model, leaving J(J-1)/2 - 1 estimable parameters. For a complete discussion of this issue, see Bunch (1991).

For practical estimation it is necessary to ensure that covariance matrices be positive (semi-) definite. This leads to our approach of estimating the J(J-1)/2-1 parameters for the lower triangular matrix $B_J$ given by

$$B_J = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ b_{21} & b_{22} & 0 & \cdots & 0 \\ b_{31} & b_{32} & b_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{J-1\,1} & b_{J-1\,2} & b_{J-1\,3} & \cdots & b_{J-1\,J-1} \end{bmatrix}, \tag{3}$$

where $B_J$ is the Cholesky factor of a (J-1)-order covariance matrix of error-term differences. By convention, this matrix is derived by using differences ($\varepsilon_j$ - $\varepsilon_J$), where J is the index of the last choice alternative appearing in C. In addition, if alternative-specific means are also being estimated, then $\mu_J \equiv 0$ by convention. Once the $B_J$ matrix is estimated, it is possible to compute what $\Sigma_\varepsilon$ would be under various equivalent normalizations--see Bunch (1991).

## 3. General Estimation Framework

Much of the code in MLMNP/B may be regarded as a "prototype" program (denoted "MLEPCM" for "Maximum Likelihood Estimation of Probabilistic Choice Models") which minimizes the following negative log-likelihood function:

$$- \sum_{I=1}^{NOBS} WT(I) * \log P[ICH(I), IX(I), RX(I)]$$

where:

P[ICH(I), IX(I), RX(I)] is a general probabilistic choice model,
ICH(I) is an integer indicating the choice made for observation I,
IX(I) contains integer explanatory data specific to observation I,
RX(I) contains real explanatory data specific to observation I, and
WT(I) is a weight for observation I.

For example, IX(I) might be a scalar value containing the number of choices available for observation I, or it might be an array used to store indexes identifying the specific choices available for observation I (e.g., ICH(I) = 4 and IX(I) contains "1 2 4 6") when such integers are required (see examples below). (This is how IX(I) is used for the MNP model, but alternative uses might be appropriate for other models.) RX(I) is usually an array containing "generic" explanatory data for all the available choices in observation I.

WT(I) is a multiplicative weighting factor convenient for economizing on computation; it is especially useful when there are repeated choices for observations which have identical explanatory information. For example, suppose that ten DM's having exactly the same explanatory data each make a choice from a choice set containing the same three alternatives, and suppose that the observed choice frequencies are (2, 7, 1). Then these ten observations can be written instead as three observations with weights of 2, 7, and 1, respectively. This requires that only three probabilities be computed, rather than ten.

## 4. Input Format

The basic input parameter stream is read from Fortran Unit 1. The order of input information is as follows:

1. General input parameters (related to the framework in section 3, all on one line).
2. Input parameters related to regression diagnostics.
3. Model-specific input parameters (in this case for the MNP model, all on one line).
4. Starting values for the iterative search (plus simple bounds for MLMNPB).

For ease of use, both MLMNP and MLMNPB use exactly the same input format, despite the fact that MLMNP does not enforce bounds and MLMNPB does not calculate statistics. This means that some elements of the input stream will be redundant for either program, but this avoids unnecessary duplication of effort in creating input files.

The data stream including observed choices and model explanatory variables is read from Fortran Unit 2. Repeated "blocks" of data are used, with each block containing all the data for one observation. Input format descriptions are provided here, with examples appearing in section 5.

### 4.1 General input parameters and regression diagnostic parameters.

The general input parameters are entered on one line, using free format as follows:

```
C  *** GENERAL PARAMETERS ARE READ IN FIRST FROM "INPUT BLOCK 1": ***
C
C    READ(1,*) NPAR,NOBS,ICSET,WEIGHT,NIVAR,NRVAR,IOUNIT,IPRNT,COVTYP,IDR
```

If IDR < 3, then regression diagnostics are of the straightforward "leave-one-out" variety. If IDR = 3, then the regression diagnostics are of the "leave-block-out" variety, and additional lines are required. Excerpts of program comments explaining the usage of all these input parameters, including the additional lines for regression diagnostics are reproduced here.

```
C  ***  MLEPCM PARAMETER USAGE ***
C
C (SEE EXPLANATIONS BELOW)
C
C SCALARS:
C
C BS...... BLOCK-SIZE, IF LEAVE-BLOCK-OUT REGRESSION DIAGNOSTICS ARE
C             REQUESTED AND ALL BLOCKS ARE THE SAME SIZE (SEE BELOW).
C COVTYP.. INDICATES TYPE OF VARIANCE-COVARIANCE MATRIX APPROXIMATION.
C             = 1 FOR H^-1, WHERE H IS THE FINITE-DIFFERENCE HESSIAN
C                 AT THE SOLUTION.
C             = 2 FOR (J^T J)^-1, I.E., THE GAUSS-NEWTON HESSIAN
C               APPROXIMATION AT THE SOLUTION.
C ICSET... INDICATOR OF FIXED- OR VARIABLE-SIZE CHOICE SETS.
C IDR..... INDICATOR FOR TYPE OF REGRESSION DIAGNOSTICS (SEE BELOW).
C IOUNIT.. OUTPUT UNIT NUMBER FOR PRINTING ERROR MESSAGES.
C              = FORTRAN UNIT FOR IOUNIT > 0.  DEFAULT = 6.
C IPRNT... INDEX INDICATING PRINT OPTIONS.
C              = 0 FOR NO ADDITIONAL PRINTING.
C              = 1 FOR FINAL CHOICE PROBABILITIES.
C              (DEFAULT = 0.)
C WEIGHT.. INDICATOR FOR USER-PROVIDED WEIGHTS.
C NB...... NUMBER OF BLOCKS, IF LEAVE-BLOCK-OUT REGRESSION DIAGNOSTICS
C             ARE REQUESTED (SEE BELOW).
C NFIX.... PARAMETER USED BY DRGLG.  NFIX = 0.
C NIVAR... NUMBER OF (INTEGER) DATA VARIABLES PER CHOICE SET.
C NIUSER.. NUMBER OF (INTEGER) USER-SPECIFIED CONSTANTS.
C NOBS.... NUMBER OF OBSERVATIONS.
C NPAR.... NUMBER OF MODEL PARAMETERS (X COMPONENTS).
```

```
C NRVAR... NUMBER OF (REAL) DATA VARIABLES PER CHOICE SET.
C NRUSER.. NUMBER OF (REAL) USER-SPECIFIED CONSTANTS.
C XNOTI... INDICATOR FOR TYPE OF REGRESSION DIAGNOSTICS (SEE BELOW).
C
C  ***  DISCUSSION FOR MLEPCM ***
C
C  ***  DATA INPUT STREAM ***
C
C  *** GENERAL PARAMETERS ARE READ IN FIRST FROM "INPUT BLOCK 1": ***
C
C    READ(1,*) NPAR,NOBS,ICSET,WEIGHT,NIVAR,NRVAR,IOUNIT,IPRNT,COVTYP,IDR
C
C      THESE PARAMETERS ARE INTENDED TO GIVE A FLEXIBLE INPUT
C   FORMAT FOR CHOICE MODELS, WITH SOME SHORTCUTS FOR SIMPLE CASES.
C   SPECIFIC SETTINGS OF THE ABOVE PARAMETERS WILL PRODUCE DIFFERENCES
C   IN THE INPUT STREAM FORMAT.
C
C   FOR ICSET = 0 (OR 1) A VARIABLE NUMBER OF ALTERNATIVES PER CHOICE
C      SET IS USED.  THE USER MUST PROVIDE THIS NUMBER FOR EACH
C      OBSERVATION.
C   FOR ICSET > 1 EACH CHOICE SET IS ASSUMED TO INCLUDE ICSET
C      ALTERNATIVES.
C
C   WEIGHT = 1 MEANS THAT EACH OBSERVATION REQUIRES A WEIGHT, WHICH
C      MUST BE PROVIDED BY THE USER.
C   WEIGHT = 0 MEANS THAT ALL OBSERVATIONS AUTOMATICALLY RECEIVE EQUAL
C      WEIGHT AND THEREFORE NO USER-SUPPLIED WEIGHTS ARE REQUIRED.
C
C   FOR NIVAR = -1 NO INTEGER DATA VALUES ARE REQUIRED BY THE MODEL.
C   FOR NIVAR =  0 A VARIABLE NUMBER OF INTEGER DATA VALUES IS STORED
C      PER OBSERVATION.  IN THIS CASE, THE USER MUST INCLUDE FOR EACH
C      OBSERVATION THE NUMBER OF INTEGER VALUES TO BE STORED FOLLOWED
C      BY THE INTEGER VALUES THEMSELVES.  (THIS MIGHT BE USED IN
C      CONJUNCTION WITH ICSET=0 TO LIST NOMINAL VARIABLES FOR THE
C      CHOICE ALTERNATIVES IN THE CHOICE SET.)
C   FOR NIVAR > 0 EACH OBSERVATION IS ASSUMED TO INCLUDE NIVAR INTEGERS.
C
C   FOR NRVAR THE USAGE IS ANALOGOUS TO NIVAR, ONLY FOR REAL DATA.
C
C   NIUSER AND NRUSER ARE USED TO INDICATE THE NUMBER OF CONSTANTS
C      TO BE PASSED TO THE MODEL SUBROUTINES.  THESE ARE MODEL SPECIFIC.
C      FOR SOME CODES NIUSER, NRUSER, AND PERHAPS THE CONSTANTS, MIGHT
C      BE SET IN THE MAIN PROGRAM AND NOT BY THE INPUT STREAM.
C
C   FOR MORE DETAILS ON THIS, SEE THE ACTUAL CODE BELOW.
C
C     IN ADDITION TO DATA STORAGE, MLEPCM PROVIDES A RATHER FLEXIBLE
C   CHOICE OF STATISTICAL ANALYSES.  IN THE VERSION OF THE PROGRAM
C   WHICH ENFORCES BOUNDS, STATISTICS ARE NOT CALCULATED.  HOWEVER,
C   FOR CONVENIENCE IT IS ASSUMED THAT THE SAME INPUT STREAM IS USED
C   FOR BOTH PROGRAMS.
C
C      TO CALCULATE ASYMPTOTIC T-SCORES, A VARIANCE-COVARIANCE MATRIX
C   APPROXIMATION IS REQUIRED.  SEE COVTYP ABOVE.
C
C      TO PERFORM REGRESSION DIAGNOSTICS, THE FOLLOWING PARAMETERS
C   ARE USED:
C
```

```
C   IDR = 0 IF NO REGRESSION DIAGNOSTICS ARE DESIRED.
C
C       = 1 FOR ONE-STEP ESTIMATES OF F(X*)-F(X(I)), WHERE X(I)
C            MINIMIZES F (THE NEGATIVE LOG-LIKELIHOOD) WITH
C            OBSERVATION I REMOVED, AND X* IS THE MLE FOR THE FULL
C            DATASET. ("LEAVE-ONE-OUT" DIAGNOSTICS.)
C
C       = 2 FOR ONE-STEP ESTIMATES OF F(X*)-F(X(I)) AS WHEN IDR = 1,
C            AND ALSO THE ONE-STEP ESTIMATES OF X(I), I = 1 TO NOBS.
C
C       = 3 FOR "LEAVE-BLOCK-OUT" DIAGNOSTICS.  (DISCUSSION FOLLOWS.)
C
C  *** PARAMETERS RELATED TO "LEAVE-BLOCK-OUT" REGRESSION DIAGNOSTICS ***
C  *** READ NEXT FROM "INPUT BLOCK 2" (IF APPLICABLE).                ***
C
C   "LEAVE-BLOCK-OUT" DIAGNOSTICS
C
C       IN THIS CASE, ONE OR MORE ADDITIONAL LINES OF DATA ARE
C    REQUIRED. IF IDR = 3, THE FOLLOWING STATEMENT IS EXECUTED:
C
C             READ(1,*) BS, NB, XNOTI
C
C    NB = NUMBER OF BLOCKS
C
C    XNOTI = 0 IF NO X(I) DIAGNOSTICS ARE REQUESTED,
C          = 1 OTHERWISE.
C
C    BS > 0 MEANS THAT FIXED BLOCK SIZES OF SIZE BS ARE USED.
C           IN THIS CASE NB * BS = NOBS, AND THE PROGRAM
C           PROCEEDS TO "INPUT BLOCK 3" FOR MNP INPUT PARAMETERS.
C
C    BS = 0 MEANS THAT VARIABLE BLOCK SIZES ARE USED.
C           IN THIS CASE THE NEXT FORMAT STATEMENT READS
C           THE BLOCK SIZES INTO RHOI USING FREE FORMAT:
C
C           LR1 = LUI + 1
C           LR2 = LR1 + NB
C           READ(1,*) (RHOI(I),I=LR1,LR2)
C
C  *** THE PROGRAM THEN PROCEEDS TO "INPUT BLOCK 3" TO READ MODEL-***
C  *** RELATED PARAMETERS.  SEE DISCUSSION FOR MNP MODEL BELOW.   ***
```

## 4.2 Multinomial probit model input parameters

After the above one (or two) "input blocks" have been processed, the next line contains five integer values which relate specifically to the linear-in-parameters MNP model described in section 2. The relevant lines from the program, plus comments are:

```
C *** READ MNP PARAMETERS FROM "INPUT BLOCK 3" ***
      READ(1,*) NALT, NATTR, IDUM, ICOV, ITASTE
C
C *** MULTINOMIAL PROBIT MODEL PARAMETERS ***
C THE FOLLOWING VARIABLES ARE USER-PROVIDED INTEGER CONSTANTS:
C IDUM.... INDICATOR FOR ALTERNATIVE-SPECIFIC MEANS,
C              = 0 FOR NO, = 1 FOR YES.  IF ICSET .NE. 0, THEN
C              THE SAME SET OF MEANS IS USED FOR EACH CHOICE SET.
C              OTHERWISE, INTEGER DATA SHOULD BE USED TO IDENTIFY THE
```

```
C              ALTERNATIVES IN EACH CHOICE SET (SEE NALT BELOW).
C ICOV.... INDICATOR FOR TYPE OF ALTERNATIVE-SPECIFIC ERRORS,
C              = 0 FOR IID ERRORS, = 1 FOR CORRELATED ERRORS.
C              IF ICSET .NE. 0, THEN THE SAME CORRELATION MATRIX IS
C              USED FOR EVERY SUBSET.  OTHERWISE, INTEGER DATA SHOULD
C              BE USED TO IDENTIFY THE ALTERNATIVES IN EACH CHOICE SET.
C ITASTE.. INDICATOR FOR TASTE VARIATION,
C              = 0 FOR NO TASTE VARIATION, = 1 FOR UNCORRELATED TASTE
C              VARIATION, = 2 FOR CORRELATED TASTE VARIATION.
C NALT.... TOTAL NUMBER OF NOMINAL CHOICE ALTERNATIVES (IF APPLICABLE).
C              IF ICSET .NE. 0, THEN NALT IS SET EQUAL TO ICSET.
C              OTHERWISE, NALT SHOULD BE > 0 IF EITHER IDUM OR ICOV
C              (OR BOTH) ARE > 0.
C NATTR... NUMBER OF ATTRIBUTES (I.E., REAL DATA VARS.) PER ALTERNATIVE.
```

A model only requires nominal variables if it (i) includes alternative-specific means, or (ii) uses a general covariance matrix, or (iii) both. This may occur with either fixed or flexible choice sets.

### 4.3 Starting values for iterative search

The following code is used to read in information for the model parameters (stored in the array X) from Fortran Unit 1:

```
      DO 10 I = 1, NPAR
          READ(1,3) VNAME(I)
    3     FORMAT(1X,A8)
          READ(1,*) X(I), B(1,I), B(2,I)
            WRITE(IOUNIT,4) I, VNAME(I),X(I), B(1,I), B(2,I)
    4   FORMAT(1X,I2,1X,A8,2X,3(1X,D13.6))
   10 CONTINUE
```

In addition to the initial starting value for X(I) in the read statement, B(1, I) contains the lower bound for parameter X(I), B(2, I) contains the upper bound for parameter X(I), and VNAME(I) contains an eight-character name for the parameter X(I).

"Blocks" of model parameters must appear in the following order, where the number of each type of parameter is indicated in brackets [].

1. Coefficients for generic attributes [NATTR]. No input if NATTR=0.
2. Coefficients for alternative-specific means [NALT-1]. No input if IDUM = 0.
3. Covariance parameters for alternative-specific errors, i.e., for Cholesky factor $B_J$ (see equation 1.3) [LCOVX = NALT(NALT-1)/2 - 1]. No input if ICOV = 0.
4. Parameters for random taste variation. No input if NATTR=0.
   No input if ITASTE=0 (no taste variation).
   NATTR parameters if ITASTE=1 (uncorrelated taste => diagonal covariance matrix).
   NATTR(NATTR+1)/2 if ITASTE=2 (correlated taste => Cholesky factor of $\Sigma_\delta$).

The number of model parameters, NPAR, is automatically checked for accuracy. If NPAR is inconsistent with the other basic input parameters, an error message is printed and the program stops. Recall that MLMNP reads in values for the bounds, but does not use them.

### 4.4 Repeated blocks of choice and explanatory data

Choice and explanatory data are read in from Fortran Unit 2. This part of the input stream is the most complex, since the actual format will vary depending upon the basic input parameter values assigned previously. One "block" of data is read for each observation. Each block is divided into "sub-blocks," where a sub-block contains a particular type of data. In general, data are read using free format. The sub-blocks are read in as follows:

<u>Sub-block 1.</u> Choice information (required).
ICH(I) is the index of the observed choice; if an inappropriate integer is entered, an error message is produced. NCH(I) is the number of choices available.
Format(s):

```
IF (ICSET.EQ.0) THEN
   READ(2,*) ICH(I), NCH(I)
ELSE
   NCH(I) = ICSET
   READ(2,*) ICH(I)
ENDIF
```

<u>Sub-block 2.</u> Integer explanatory data (no input if NIVAR = -1).
Let IX(I,*) be an array including explanatory data for observation I, and LIX(I) be the length of IX(I,*). For this program, IX(I,*) contains integer values indicating the possible choices available in the choice set. This sub-block is only input when NIVAR ≥ 0, but is required for MNP when the following two conditions hold: (i) flexible choice sets are used (ICSET = 0), and (ii) either IDUM or ICOV (or both) are > 0. Note: LIX(I) should be equal to NCH(I) in these cases; although this is redundant, it preserves the flexibility of the input format for other models. If flexible choice sets are used and the size of the choice set may vary, then NIVAR = 0 and the first integer in the sub-block should be LIX(I). If flexible choice sets are used, but the size of the choice set stays the same for each observation, then NIVAR = NCH(I) > 0 for all I. In this case NIVAR integers are input automatically and LIX(I) is not required; the first value in the sub-block is IX(I,1). If anything other than an integer in the range 1 to NALT is used in the sub-block, an error message is produced.

Format(s):

```
IF (NIVAR.EQ.0) THEN
   READ(2,*) LIX(I), (IX(I,K),K=1,LIX(I))
ENDIF
IF (NIVAR.GT.0) THEN
   LIX(I) = NIVAR
   READ(2,*) (IX(I,K),K=1,LIX(I))
ENDIF
```

Sub-block 3. Weights (no input if WEIGHT = 0).

Weights are required only if WEIGHT = 1; one weight appears on a line.

Format(s):

```
IF (WEIGHT.EQ.1) THEN
   READ(2,*) WT(I)
ELSE
   WT(I) = ONE
ENDIF
```

Sub-block 4. Real explanatory data (no input if NRVAR = -1).

Let RX(I,*) be an array including explanatory data for observation I, and LRX(I) be the length of RX(I,*). For this program, RX(I,*) should contain NATTR values per choice alternative, i.e., LRX(I) = NATTR*NCH(I). The first NATTR attributes correspond to alternative 1, the second NATTR attributes correspond to alternative 2, etc.

This sub-block is only input when NRVAR ≥ 0. If flexible choice sets are used, then NRVAR = 0 and LRX(I) is the first value entered in the sub-block. Note: LRX(I) should equal NATTR*NCH(I) and is technically redundant, but preserves the flexibility of the input format for other models. If fixed choice sets are used, then NRVAR > 0 and NRVAR values are read for each observation; in this case LRX(I) is not required, and the first value appearing in the sub-block is RX(I,1). (NRVAR should equal NATTR*ICSET in this case.)

Format(s):

```
IF (NRVAR.EQ.0) THEN
   READ(2,*) LRX(I), (RX(I,K),K=1,LRX(I))
ENDIF
IF (NRVAR.GT.0) THEN
   LRX(I) = NRVAR
   READ(2,*) (RX(I,K),K=1,LRX(I))
ENDIF
```

## 5. Examples

This section gives examples of various types of multinomial probit models that fit the framework described in section 2, and uses them to illustrate usage of MLMNP/B.

### 5.1 Transportation mode choice

Choice models have been widely used in transportation demand analysis, particularly for choice of transport mode for a shopping trip or the commute to work. A hypothetical example data set for a three-alternative modal-split problem is given in Daganzo (1979, pg. 61). The choice set C for this example is {bus, streetcar, automobile}, and the attractiveness of mode i is assumed to be proportional to the negative of the travel time t. We use this data set to construct two examples.

Example 1. Mode choice with no taste variation.

We now describe the relevant features of this problem. First consider the general input parameters. Choices are observed for 50 decision makers, each is assumed to choose from the same fixed choice set {bus, streetcar, automobile}, travel times are different for each decision maker, but are assumed to be generic (i.e., the utility for travel time is the same regardless of the mode). This gives:

NPAR = ? (to be calculated below)
NOBS = 50
ICSET = 3 (fixed choice set of size three)
WEIGHT = 0 (equal weight for each observation: no user-supplied weights required)
NIVAR = -1 (fixed choice sets, therefore no integer data required)
NRVAR = 3 (fixed choice sets of size three, with one generic attribute per alternative)
IOUNIT = 6 (standard unit for Fortran output)
IPRNT = 0 (no printing of final choice probabilities)
COVTYP = 2 (use Gauss-Newton Hessian for covariance matrix approximation)
IDR = 0 (no regression diagnostics requested)

Next we consider MNP input parameters. There are three alternatives per choice set, and one generic attribute per alternative, as previously noted. There are many missing attributes which might determine mode choice other than travel time, and we therefore expect that: (i) alternative-specific means might be significant, and (ii) alternative-specific errors might be correlated (especially bus versus streetcar). For the time being we assume that there is no random taste variation. This gives:

NALT = 3
NATTR = 1
IDUM = 1
ICOV = 1
ITASTE = 0

Next we consider the model parameters required for this specification. There is one coefficient for travel time, two (= NALT-1) alternative-specific means for {bus, streetcar} ($\mu_{auto} \equiv 0$ by convention), two (= NALT(NALT-1)/2-1) alternative-specific error covariance parameters, and no taste variation covariance parameters, giving a total of five parameters to be estimated (NPAR = 5). They appear in the following order:

TTIME
DBUS
DSTREETC
B21
B22

Finally, there is the input format for the choices and explanatory data. For this particular set of basic input parameters, the block for observation I would consist of only sub-block 1 and sub-block 4 (see section 4) with the following format:

Input 1: Choice_index (1 for bus, 2 for streetcar, 3 for automobile)

Input 2: travel_time_1(for bus) travel_time_2(for streetcar) travel_time_3(for automobile)

The complete Fortran Unit 1 input stream for Example 1 is given below (all left-most entries start in column 2); it is available in file MNPEX1.FU1. An excerpt from the full 50-observation input stream for Fortran Unit 2 is also given below; it is available in the file DAGANZO.FU2. Finally, the full MLMNP output for this example is given below. The code produces asymptotic t-statistics using the Gauss-Newton Hessian to approximate the covariance matrix; no regression diagnostics were requested. Note that, because there are three nominal choices in every choice set (ICSET=3), it makes sense to compute the log-likelihood for a simple model using alternative-specific constants only for comparison purposes.

Fortran Unit 1 Input Stream for Example 1: File MNPEX1.FU1

```
5 50 3 0 -1 3 6 0 2 0
3 1 1 1 0
TTIME
  0.  -100.  100.
DBUS
  0.  -100.  100.
DSTREETC
  0.  -100.  100.
B21
  1.  -100.  100.
B22
  1.  -100.  100.

Example 1:  Trinomial probit model for mode choice, using Daganzo data.
  This example includes alternative-specific means, correlated alternative-
  specific errors, and NO random taste variation for traveltime.
  Explanatory data (for Unit 2 input stream) are in Daganzo.fu2.
```

Fortran Unit 2 Input Stream for Examples 1 and 2: Excerpt from File DAGANZO.FU2

```
2
16.481  16.196  23.890
2
15.123  11.373  14.182
2
19.469  8.822   20.819
2
18.847  15.649  21.280
2
12.578  10.671  18.335
1
11.513  20.582  27.838
1
10.651  15.537  17.418
1
8.359  15.675  21.050
1
11.679  12.668  23.104
2
23.237  10.356  21.346
3
13.236  16.019  10.087
3
20.052  16.861  14.168
2
18.917  14.764  21.564
2
18.200   6.868  19.095
1
10.777  16.554  15.938
2
20.003   6.377   9.314
2
19.768  8.523   18.960
 (etc.)
```

Output for Example 1

```
PROGRAM MLMNP

MAXIMUM LIKELIHOOD ESTIMATION OF
LINEAR-IN-PARAMETERS MULTINOMIAL PROBIT MODELS
(BOUNDS NOT ENFORCED; STATISTICS ARE COMPUTED)


 NUMBER OF OBSERVATIONS.................  50
 NUMBER OF ALTERNATIVES PER CHOICE SET..   3
 EQUAL WEIGHTS FOR ALL OBSERVATIONS
 NO INTEGER EXPLANATORY VARIABLES
 NUMBER OF REAL DATA VALUES PER OBS.....   3
 OUTPUT UNIT............................   6

 COVARIANCE TYPE = INVERSE GAUSS-NEWTON HESSIAN
 NO REGRESSION DIAGNOSTICS REQUESTED

 *** NOTE:  NALT SET EQUAL TO ICSET ***
 NUMBER OF NOMINAL VARIABLES............   3
 NUMBER OF ATTRIBUTES PER ALTERNATIVE...   1
 NOMINAL DUMMIES USED
 CORRELATED ERROR TERMS
 NO RANDOM TASTE VARIATION

 NUMBER OF MODEL PARAMETERS.............   5

INITIAL PARAMETER VECTOR AND BOUNDS:
 1 TTIME       0.000000D+00 -0.100000D+03  0.100000D+03
 2 DBUS        0.000000D+00 -0.100000D+03  0.100000D+03
 3 DSTREETC    0.000000D+00 -0.100000D+03  0.100000D+03
 4 B21         0.100000D+01 -0.100000D+03  0.100000D+03
 5 B22         0.100000D+01 -0.100000D+03  0.100000D+03

NONDEFAULT VALUES....

DTYPE..... IV(16) =  0
DINIT..... V(38) =  0.1000000D+01

    I     INITIAL X(I)        D(I)

    1     0.000000D+00     0.100D+01
    2     0.000000D+00     0.100D+01
    3     0.000000D+00     0.100D+01
    4     0.100000D+01     0.100D+01
    5     0.100000D+01     0.100D+01

   IT   NF      F        RELDF   PRELDF   RELDX  MODEL  STPPAR  D*STEP  NPRELDF

    0    1 0.546D+02
    1    2 0.390D+02 0.29D+00 0.24D+00 0.4D+00    G   0.2D+00 0.9D+00 0.46D+00
    2    3 0.341D+02 0.13D+00 0.13D+00 0.5D+00    G   0.2D+01 0.1D+01 0.26D+00
    3    4 0.331D+02 0.30D-01 0.41D-01 0.3D+00    G   0.0D+00 0.5D+00 0.41D-01
    4    5 0.324D+02 0.20D-01 0.18D-01 0.2D+00    S   0.0D+00 0.2D+00 0.18D-01
    5    6 0.323D+02 0.44D-02 0.42D-02 0.7D-01    S   0.0D+00 0.1D+00 0.42D-02
    6    7 0.323D+02 0.20D-03 0.22D-03 0.1D-01    S   0.0D+00 0.2D-01 0.22D-03
    7    8 0.323D+02 0.17D-04 0.16D-04 0.5D-02    S   0.0D+00 0.8D-02 0.16D-04
```

```
    8    9 0.323D+02 0.38D-06 0.36D-06 0.6D-03    S   0.0D+00 0.1D-02 0.36D-06
    9   10 0.323D+02 0.30D-08 0.36D-08 0.6D-04    G   0.0D+00 0.9D-04 0.36D-08
   10   11 0.323D+02 0.47D-09 0.11D-08 0.6D-04    G   0.0D+00 0.1D-03 0.11D-08
   11   12 0.323D+02 0.32D-09 0.36D-09 0.2D-04    S   0.0D+00 0.3D-04 0.36D-09
   12   13 0.323D+02 0.13D-10 0.26D-10 0.7D-05    S   0.0D+00 0.1D-04 0.26D-10

***** RELATIVE FUNCTION CONVERGENCE *****

FUNCTION      0.322697D+02   RELDX         0.682D-05
FUNC. EVALS       13         GRAD. EVALS      13
PRELDF        0.257D-10      NPRELDF       0.257D-10

    I      FINAL X(I)        D(I)          G(I)

    1   -0.113148E+00     0.100E+01    -0.749E-04
    2    0.127572E-01     0.100E+01     0.149E-04
    3    0.198355E+00     0.100E+01    -0.136E-05
    4    0.592161E+00     0.100E+01    -0.943E-04
    5    0.312716E+00     0.100E+01    -0.507E-04

   1 EXTRA FUNC. EVALS FOR COVARIANCE AND DIAGNOSTICS.
   1 EXTRA GRAD. EVALS FOR COVARIANCE AND DIAGNOSTICS.

SQRT(RECIPROCAL CONDITION OF (J**T)*RHO"*J) = AT MOST  0.48E-01

COVARIANCE = (J**T * RHO" * J)**-1

ROW  1     0.209E-02
ROW  2     0.514E-02   0.147
ROW  3     0.645E-02   0.113       0.999E-01
ROW  4    -0.241E-02   0.549E-01   0.385E-01   0.463E-01
ROW  5    -0.131E-01  -0.101E-01  -0.274E-01   0.286E-01   0.121

REGRESSION DIAGNOSTIC VECTOR NOT COMPUTED

ASYMPTOTIC T-STATISTICS:
 I               X(I)           T-STAT(I)        STD ERROR
 1  TTIME     -0.113148D+00   -0.247728D+01    0.456744D-01
 2  DBUS       0.127572D-01    0.332360D-01    0.383837D+00
 3  DSTREETC   0.198355D+00    0.627558D+00    0.316075D+00
 4  B21        0.592161D+00    0.275154D+01    0.215211D+00
 5  B22        0.312716D+00    0.898064D+00    0.348211D+00

NUMBER OF OBSERVATIONS (NOBS) =   50

LOG-LIKELIHOOD L(EST)  = -0.322697D+02
LOG-LIKELIHOOD L(0)    = -0.549306D+02
-2[L(0) - L(EST)]:     =  0.453219D+02

1 - L(EST)/L(0):       =  0.412538D+00
1 - (L(EST)-NPAR)/L(0) =  0.321514D+00

 (FIXED CHOICE SET SIZE)

AGGREGATE CHOICES AND MARKET SHARES:
  1      14.000  0.2800
  2      29.000  0.5800
  3       7.000  0.1400
```

```
STATISTICS FOR CONSTANTS-ONLY MODEL:
   LOG-LIKELIHOOD L(C)    = -0.473814D+02
   -2[L(C) - L(EST)]:     =  0.302235D+02


OUTPUT FOR CONVENIENT RESTART:
TTIME
 -0.113148D+00 -0.100000D+03  0.100000D+03
DBUS
  0.127572D-01 -0.100000D+03  0.100000D+03
DSTREETC
  0.198355D+00 -0.100000D+03  0.100000D+03
B21
  0.592161D+00 -0.100000D+03  0.100000D+03
B22
  0.312716D+00 -0.100000D+03  0.100000D+03
```

(End of output for Example 1.)

Example 2. Mode choice with taste variation.

All the details for this example are the same as for Example 1, except that we assume that there is random taste variation for travel time. The only require change is to set ITASTE = 1 for uncorrelated random taste variation. Note: Since there is only one generic attribute for this case, setting ITASTE = 2 produces the same results. The relevant Fortran Unit 1 file follows:

Fortran Unit 1 Input Stream for Example 2: File MNPEX2.FU1

```
6 50 3 0 -1 3 6 0 2 0
3 1 1 1 1
TTIME
  0.  -100.  100.
DBUS
  0.  -100.  100.
DSTREETC
  0.  -100.  100.
B21
  1.  -100.  100.
B22
  1.  -100.  100.
SigT
  1.   0.0001  100.
Example 2:  Trinomial probit model for mode choice, using Daganzo data.
  This example includes alternative-specific means, correlated alternative-
  specific errors, and random taste variation for traveltime.
  Explanatory data (for Fortran Unit 2) are in Daganzo.fu2.
```

## 5.2 Choice of rental dwelling by MBA students.

Choice models are also widely used in marketing research. A prototypical example used in a number of research studies involve predicting the choice of a place to live by MBA students. In a recent study conducted by students in the Graduate School of Management, UC Davis, survey data was collected using questions of the following type:

```
*******************************************************************************
1.   Suppose you're looking for housing in Davis for next year, and you're leaving your
     current place.  The following rentals are available.  Which would you choose?  Please
     circle the letter corresponding to your selection.

     Rent     Location                Condition             Bedrms  Housing Type

A:   $900   North/West/East    New, or refurbished as new     2    Duplex/house

B:   $650   Central Davis      Needs maintenance              2    Apartment

C:   Residence similar to your current place.
*******************************************************************************
```

An experimental design was used to vary rent, location, condition, and number of bedrooms using three levels each. Housing type had two levels. An interesting feature of this study is that each participant answered 27 questions. The final sample contained 21 individuals, giving a total of 567 choices. Characteristics of choice "C" were collected in a separate part of the survey. For a data set of this type, leave-block-out regression diagnostics are of interest. In other words, data may be left out one individual at a time by specifying a fixed block-size of 27, and setting the number of blocks to 21. Two more examples with printouts are included below. The Fortran Unit 2 file is RENT.FU2.

<u>Example 3.</u> Dwelling choice and fixed block-size regression diagnostics.
The following input stream specifies a simple IID probit model with a quantitative attribute for rent, and two dummy variables for each of location, condition, and number of bedrooms, respectively. In addition, a single dummy variable is used for housing-type. Finally, a dummy variable is used to examine the the alternative-specific effect for choice "C."

Fortran Unit 1 Input Stream for Example 3: File RENT1.FU1

```
 9 567 3 0 -1 27 6 0 2 3
27 21 1
3 9 0 0 0
RENT
 -0.371499D-02 -0.100000D+03  0.100000D+03
LocD1
  0.473069D-01 -0.100000D+03  0.100000D+03
LocD2
 -0.443496D+00 -0.100000D+03  0.100000D+03
ConD1
  0.734521D+00 -0.100000D+03  0.100000D+03
ConD2
  0.648764D+00 -0.100000D+03  0.100000D+03
BedD1
 -0.125812D+01 -0.100000D+03  0.100000D+03
BedD2
 -0.641347D+00 -0.100000D+03  0.100000D+03
Htype
  0.429202D+00 -0.100000D+03  0.100000D+03
CDum
 +0.958062D+00 -0.100000D+03  0.100000D+03

Example 3:  Simple trinomial probit model with no exploitation of
demographics.  (This also works with trinomial probit)
Includes an alternative specific dummy for alternative "C"
  Explanatory data (for Unit 2 input stream) are in RENT.FU2

Gauss-Newton Hessian
Leave-Block-Out Regression Diagnostics
  Block size is fixed, = 27
  No. of blocks = 21
  X(I) diagnostics are requested
```

Output for Example 3

```
PROGRAM MLMNP

MAXIMUM LIKELIHOOD ESTIMATION OF
LINEAR-IN-PARAMETERS MULTINOMIAL PROBIT MODELS
(BOUNDS NOT ENFORCED; STATISTICS ARE COMPUTED)


 NUMBER OF OBSERVATIONS................. 567
 NUMBER OF ALTERNATIVES PER CHOICE SET..   3
 EQUAL WEIGHTS FOR ALL OBSERVATIONS
 NO INTEGER EXPLANATORY VARIABLES
 NUMBER OF REAL DATA VALUES PER OBS.....  27
 OUTPUT UNIT...........................   6

 COVARIANCE TYPE = INVERSE GAUSS-NEWTON HESSIAN
 REGRESSION DIAGNOSTICS REQUESTED

 *** LEAVE-BLOCK-OUT DIAGNOSTICS REQUESTED  ***

 DIAGNOSTICS ON X-VECTOR REQUESTED
 NUMBER OF BLOCKS:    21
 FIXED BLOCK SIZE:    27
```

```
 *** NOTE:  NALT SET EQUAL TO ICSET ***
 NUMBER OF NOMINAL VARIABLES...........   3
 NUMBER OF ATTRIBUTES PER ALTERNATIVE...   9
 NO NOMINAL DUMMIES
 IID ERROR TERMS
 NO RANDOM TASTE VARIATION

 NUMBER OF MODEL PARAMETERS............   9

INITIAL PARAMETER VECTOR AND BOUNDS:
 1 RENT       -0.371499D-02 -0.100000D+03  0.100000D+03
 2 LocD1       0.473069D-01 -0.100000D+03  0.100000D+03
 3 LocD2      -0.443496D+00 -0.100000D+03  0.100000D+03
 4 ConD1       0.734521D+00 -0.100000D+03  0.100000D+03
 5 ConD2       0.648764D+00 -0.100000D+03  0.100000D+03
 6 BedD1      -0.125812D+01 -0.100000D+03  0.100000D+03
 7 BedD2      -0.641347D+00 -0.100000D+03  0.100000D+03
 8 Htype       0.429202D+00 -0.100000D+03  0.100000D+03
 9 CDum        0.958062D+00 -0.100000D+03  0.100000D+03

NONDEFAULT VALUES....

DTYPE..... IV(16) =  0
DINIT..... V(38) =  0.1000000D+01

    I     INITIAL X(I)        D(I)

    1    -0.371499D-02     0.100D+01
    2     0.473069D-01     0.100D+01
    3    -0.443496D+00     0.100D+01
    4     0.734521D+00     0.100D+01
    5     0.648764D+00     0.100D+01
    6    -0.125812D+01     0.100D+01
    7    -0.641347D+00     0.100D+01
    8     0.429202D+00     0.100D+01
    9     0.958062D+00     0.100D+01

   IT   NF      F       RELDF   PRELDF   RELDX  MODEL  STPPAR  D*STEP  NPRELDF

    0    1 0.418D+03
    1    2 0.415D+03 0.65D-02 0.84D-02 0.1D+00    G   0.0D+00 0.4D+00 0.84D-02
    2    3 0.415D+03 0.35D-03 0.41D-03 0.2D-01    G   0.0D+00 0.8D-01 0.41D-03
    3    4 0.415D+03 0.15D-04 0.15D-04 0.3D-02    S   0.0D+00 0.1D-01 0.15D-04
    4    5 0.415D+03 0.14D-06 0.15D-06 0.3D-03    S   0.0D+00 0.1D-02 0.15D-06
    5    6 0.415D+03 0.22D-08 0.22D-08 0.3D-04    S   0.0D+00 0.2D-03 0.22D-08
    6    7 0.415D+03 0.42D-10 0.40D-10 0.5D-05    S   0.0D+00 0.2D-04 0.40D-10

***** RELATIVE FUNCTION CONVERGENCE *****

FUNCTION     0.414996D+03   RELDX        0.467D-05
FUNC. EVALS       7         GRAD. EVALS       7
PRELDF       0.404D-10      NPRELDF      0.404D-10

    I      FINAL X(I)        D(I)          G(I)

    1   -0.417176E-02     0.100E+01     0.679E-02
    2    0.309478E-02     0.100E+01     0.147E-03
```

```
     3   -0.415382E+00     0.100E+01    -0.615E-04
     4    0.805209E+00     0.100E+01     0.610E-05
     5    0.739921E+00     0.100E+01     0.103E-03
     6   -0.156055E+01     0.100E+01     0.931E-04
     7   -0.704689E+00     0.100E+01    -0.197E-03
     8    0.541099E+00     0.100E+01    -0.277E-03
     9    0.102279E+01     0.100E+01     0.238E-03

   1 EXTRA FUNC. EVALS FOR COVARIANCE AND DIAGNOSTICS.
   1 EXTRA GRAD. EVALS FOR COVARIANCE AND DIAGNOSTICS.

SQRT(RECIPROCAL CONDITION OF (J**T)*RHO"*J) = AT MOST  0.15E-02

COVARIANCE = (J**T * RHO" * J)**-1

ROW  1      0.122E-06
ROW  2      0.637E-05   0.139E-01
ROW  3      0.774E-05   0.743E-02   0.206E-01
ROW  4     -0.484E-05   0.946E-03  -0.166E-02   0.197E-01
ROW  5     -0.615E-05   0.153E-03  -0.299E-02   0.107E-01   0.183E-01
ROW  6      0.189E-04  -0.847E-03  -0.108E-02  -0.475E-02  -0.601E-02
            0.229E-01
ROW  7      0.906E-05   0.105E-02   0.159E-02  -0.912E-03  -0.111E-02
            0.638E-02   0.126E-01
ROW  8     -0.187E-04   0.155E-02  -0.126E-02   0.220E-02   0.379E-02
           -0.303E-02  -0.224E-02   0.156E-01
ROW  9     -0.121E-04  -0.396E-03   0.101E-03   0.321E-02   0.449E-03
           -0.396E-02  -0.127E-02   0.622E-02   0.930E-02

BLOCK REGRESSION DIAGNOSTIC = 0.5 * G(I)**T * H(I)**-1 * H * H(I)**-1 * G(I)
BLOCK  FIRST  LAST    RD(I)         X(I)
    1      1    27    1.37      -0.432037E-02   0.688989E-02  -0.454742
                                 0.750261       0.706886       -1.40656
                                -0.688482       0.532909        1.05724
    2     28    54    1.35      -0.395289E-02   0.671259E-01  -0.401830
                                 0.934180       0.764509       -1.49978
                                -0.607417       0.491203        1.05807
    3     55    81    2.28      -0.398184E-02   0.891926E-01  -0.432253
                                 0.718804       0.592961       -1.58360
                                -0.629194       0.464619        1.07271
    4     82   108   0.782      -0.418353E-02  -0.479049E-01  -0.443015
                                 0.860322       0.829490       -1.48508
                                -0.641022       0.527860       0.977591
    5    109   135   0.392      -0.416660E-02  -0.198890E-01  -0.473270
                                 0.723134       0.657318       -1.52645
                                -0.718463       0.549876        1.03077
    6    136   162   0.733      -0.422016E-02   0.369009E-01  -0.370415
                                 0.864933       0.794825       -1.49063
                                -0.624312       0.541274       0.988646
    7    163   189   0.574      -0.401477E-02   0.216230E-01  -0.333111
                                 0.844806       0.787417       -1.52104
                                -0.683333       0.569935        1.05228
    8    190   216   0.205      -0.412189E-02   0.302116E-02  -0.392089
                                 0.756986       0.671126       -1.50307
                                -0.708418       0.505785       0.999089
    9    217   243    2.02      -0.420589E-02  -0.531855E-01  -0.384752
                                 0.854976       0.722049       -1.72319
                                -0.710096       0.594932        1.00717
```

| | | | | | | |
|---|---|---|---|---|---|---|
| 10 | 244 | 270 | 0.517 | -0.413471E-02 | -0.811884E-01 | -0.426687 |
| | | | | 0.812327 | 0.754296 | -1.61269 |
| | | | | -0.710099 | 0.537952 | 1.00783 |
| 11 | 271 | 297 | 0.397 | -0.426570E-02 | -0.745965E-01 | -0.476782 |
| | | | | 0.828184 | 0.772109 | -1.52991 |
| | | | | -0.708157 | 0.589738 | 1.05453 |
| 12 | 298 | 324 | 0.793 | -0.415671E-02 | 0.640464E-01 | -0.342630 |
| | | | | 0.803615 | 0.676671 | -1.55464 |
| | | | | -0.767175 | 0.580018 | 0.998304 |
| 13 | 325 | 351 | 0.179 | -0.407197E-02 | 0.137637E-01 | -0.424389 |
| | | | | 0.754705 | 0.720375 | -1.50624 |
| | | | | -0.664943 | 0.526233 | 1.01178 |
| 14 | 352 | 378 | 23.2 | -0.520847E-02 | 0.213824 | -0.296681 |
| | | | | 0.977017 | 1.08571 | -2.15518 |
| | | | | -1.10519 | 0.589034 | 1.32706 |
| 15 | 379 | 405 | 0.303 | -0.412018E-02 | 0.369610E-02 | -0.379747 |
| | | | | 0.749824 | 0.705542 | -1.51105 |
| | | | | -0.698940 | 0.469147 | 0.981475 |
| 16 | 406 | 432 | 1.03 | -0.417480E-02 | -0.789860E-01 | -0.454166 |
| | | | | 0.883726 | 0.836766 | -1.52686 |
| | | | | -0.704554 | 0.467903 | 0.953507 |
| 17 | 433 | 459 | 2.42 | -0.426904E-02 | 0.919137E-01 | -0.269440 |
| | | | | 0.721690 | 0.799428 | -1.74527 |
| | | | | -0.692543 | 0.635327 | 1.00821 |
| 18 | 460 | 486 | 0.295 | -0.408598E-02 | -0.236992E-01 | -0.455795 |
| | | | | 0.750614 | 0.708528 | -1.52054 |
| | | | | -0.739287 | 0.533723 | 1.02839 |
| 19 | 487 | 513 | 8.07 | -0.441734E-02 | 0.487603E-01 | -0.774982 |
| | | | | 0.889980 | 0.779335 | -1.83375 |
| | | | | -0.768117 | 0.681129 | 0.998396 |
| 20 | 514 | 540 | 0.445 | -0.439653E-02 | -0.219165E-01 | -0.375854 |
| | | | | 0.822584 | 0.734316 | -1.56971 |
| | | | | -0.726309 | 0.532129 | 1.04606 |
| 21 | 541 | 567 | 0.502 | -0.408979E-02 | -0.340178E-01 | -0.413136 |
| | | | | 0.796396 | 0.693249 | -1.52454 |
| | | | | -0.755420 | 0.547292 | 0.989836 |

ASYMPTOTIC T-STATISTICS:

| I | | X(I) | T-STAT(I) | STD ERROR |
|---|---|---|---|---|
| 1 | RENT | -0.417176D-02 | -0.119269D+02 | 0.349776D-03 |
| 2 | LocD1 | 0.309478D-02 | 0.262598D-01 | 0.117853D+00 |
| 3 | LocD2 | -0.415382D+00 | -0.289368D+01 | 0.143548D+00 |
| 4 | ConD1 | 0.805209D+00 | 0.573979D+01 | 0.140286D+00 |
| 5 | ConD2 | 0.739921D+00 | 0.546424D+01 | 0.135412D+00 |
| 6 | BedD1 | -0.156055D+01 | -0.103182D+02 | 0.151244D+00 |
| 7 | BedD2 | -0.704689D+00 | -0.628199D+01 | 0.112176D+00 |
| 8 | Htype | 0.541099D+00 | 0.433084D+01 | 0.124941D+00 |
| 9 | CDum | 0.102279D+01 | 0.106073D+02 | 0.964229D-01 |

NUMBER OF OBSERVATIONS (NOBS) = 567

LOG-LIKELIHOOD L(EST) = -0.414996D+03
LOG-LIKELIHOOD L(0) = -0.622913D+03
-2[L(0) - L(EST)]: = 0.415834D+03

1 - L(EST)/L(0): = 0.333782D+00
1 - (L(EST)-NPAR)/L(0) = 0.319334D+00

```
 (FIXED CHOICE SET SIZE)

 AGGREGATE CHOICES AND MARKET SHARES:
   1     121.000  0.2134
   2     133.000  0.2346
   3     313.000  0.5520

 STATISTICS FOR CONSTANTS-ONLY MODEL:
    LOG-LIKELIHOOD L(C)    = -0.565715D+03
    -2[L(C) - L(EST)]:     =  0.301438D+03



 OUTPUT FOR CONVENIENT RESTART:
 RENT
  -0.417176D-02 -0.100000D+03  0.100000D+03
 LocD1
   0.309478D-02 -0.100000D+03  0.100000D+03
 LocD2
  -0.415382D+00 -0.100000D+03  0.100000D+03
 ConD1
   0.805209D+00 -0.100000D+03  0.100000D+03
 ConD2
   0.739921D+00 -0.100000D+03  0.100000D+03
 BedD1
  -0.156055D+01 -0.100000D+03  0.100000D+03
 BedD2
  -0.704689D+00 -0.100000D+03  0.100000D+03
 Htype
   0.541099D+00 -0.100000D+03  0.100000D+03
 CDum
   0.102279D+01 -0.100000D+03  0.100000D+03
...
```
(End of output for Example 3.)

<u>Example 4.</u> Dwelling choice and variable block-size regression diagnostics.

The model in this example is the same as for Example 3. The difference is that we wish to examine the effect of income on the preferences of the individuals in the study. The data set was sorted by income, and then blocked into three different income categories. In the three categories there were 8, 6, and 7 individuals, respectively; this corresponds to blocks of 216, 162, and 189 observations, respectively. The input stream which gives variable block-size diagnostics (including X(I) diagnostics), is as follows:

Fortran Unit 1 Input Stream for Example 4:  File RENT2.FU1

```
 9 567 3 0 -1 27 6 0 2 3
 0 3 1
 216 162 189
 3 9 0 0 0
 RENT
  -0.371499D-02 -0.100000D+03  0.100000D+03
 LocD1
   0.473069D-01 -0.100000D+03  0.100000D+03
 LocD2
  -0.443496D+00 -0.100000D+03  0.100000D+03
 ConD1
   0.734521D+00 -0.100000D+03  0.100000D+03
 ConD2
   0.648764D+00 -0.100000D+03  0.100000D+03
 BedD1
  -0.125812D+01 -0.100000D+03  0.100000D+03
 BedD2
  -0.641347D+00 -0.100000D+03  0.100000D+03
 Htype
   0.429202D+00 -0.100000D+03  0.100000D+03
 CDum
  +0.958062D+00 -0.100000D+03  0.100000D+03

 Example 4:  Simple trinomial probit model with no exploitation of
 demographics.  (This also works with trinomial probit)
 Includes an alternative specific dummy for alternative "C"
   Explanatory data (for Unit 2 input stream) are in RENT.FU2

 Gauss-Newton Hessian
 Leave-Block-Out Regression Diagnostics
   Block size is variable ( = 216, 162, 189)
   No. of blocks = 3
   X(I) diagnostics are desired...
```

Output for Example 4

```
 PROGRAM MLMNP

 MAXIMUM LIKELIHOOD ESTIMATION OF
 LINEAR-IN-PARAMETERS MULTINOMIAL PROBIT MODELS
 (BOUNDS NOT ENFORCED; STATISTICS ARE COMPUTED)


  NUMBER OF OBSERVATIONS................. 567
  NUMBER OF ALTERNATIVES PER CHOICE SET..   3
  EQUAL WEIGHTS FOR ALL OBSERVATIONS
  NO INTEGER EXPLANATORY VARIABLES
  NUMBER OF REAL DATA VALUES PER OBS.....  27
  OUTPUT UNIT...........................   6

  COVARIANCE TYPE = INVERSE GAUSS-NEWTON HESSIAN
  REGRESSION DIAGNOSTICS REQUESTED

  *** LEAVE-BLOCK-OUT DIAGNOSTICS REQUESTED  ***

  DIAGNOSTICS ON X-VECTOR REQUESTED
  NUMBER OF BLOCKS:     3
```

```
 VARIABLE BLOCK-SIZE OPTION CHOSEN

 BLOCK-SIZES:
      216  162  189

 *** NOTE:  NALT SET EQUAL TO ICSET ***
 NUMBER OF NOMINAL VARIABLES............   3
 NUMBER OF ATTRIBUTES PER ALTERNATIVE...   9
 NO NOMINAL DUMMIES
 IID ERROR TERMS
 NO RANDOM TASTE VARIATION

 NUMBER OF MODEL PARAMETERS.............   9

INITIAL PARAMETER VECTOR AND BOUNDS:
 1 RENT       -0.371499D-02 -0.100000D+03  0.100000D+03
 2 LocD1       0.473069D-01 -0.100000D+03  0.100000D+03
 3 LocD2      -0.443496D+00 -0.100000D+03  0.100000D+03
 4 ConD1       0.734521D+00 -0.100000D+03  0.100000D+03
 5 ConD2       0.648764D+00 -0.100000D+03  0.100000D+03
 6 BedD1      -0.125812D+01 -0.100000D+03  0.100000D+03
 7 BedD2      -0.641347D+00 -0.100000D+03  0.100000D+03
 8 Htype       0.429202D+00 -0.100000D+03  0.100000D+03
 9 CDum        0.958062D+00 -0.100000D+03  0.100000D+03

NONDEFAULT VALUES....

DTYPE..... IV(16) =  0
DINIT..... V(38) =  0.1000000D+01

    I     INITIAL X(I)        D(I)

    1    -0.371499D-02     0.100D+01
    2     0.473069D-01     0.100D+01
    3    -0.443496D+00     0.100D+01
    4     0.734521D+00     0.100D+01
    5     0.648764D+00     0.100D+01
    6    -0.125812D+01     0.100D+01
    7    -0.641347D+00     0.100D+01
    8     0.429202D+00     0.100D+01
    9     0.958062D+00     0.100D+01

   IT   NF      F        RELDF   PRELDF   RELDX  MODEL  STPPAR  D*STEP  NPRELDF

    0    1 0.418D+03
    1    2 0.415D+03 0.65D-02 0.84D-02 0.1D+00    G   0.0D+00 0.4D+00 0.84D-02
    2    3 0.415D+03 0.35D-03 0.41D-03 0.2D-01    G   0.0D+00 0.8D-01 0.41D-03
    3    4 0.415D+03 0.15D-04 0.15D-04 0.3D-02    S   0.0D+00 0.1D-01 0.15D-04
    4    5 0.415D+03 0.14D-06 0.15D-06 0.3D-03    S   0.0D+00 0.1D-02 0.15D-06
    5    6 0.415D+03 0.22D-08 0.22D-08 0.3D-04    S   0.0D+00 0.2D-03 0.22D-08
    6    7 0.415D+03 0.42D-10 0.40D-10 0.5D-05    S   0.0D+00 0.2D-04 0.40D-10

***** RELATIVE FUNCTION CONVERGENCE *****

FUNCTION     0.414996D+03   RELDX        0.467D-05
FUNC. EVALS       7         GRAD. EVALS       7
PRELDF       0.404D-10      NPRELDF      0.404D-10
```

```
    I      FINAL X(I)        D(I)          G(I)

    1   -0.417176E-02     0.100E+01     0.679E-02
    2    0.309478E-02     0.100E+01     0.147E-03
    3   -0.415382E+00     0.100E+01    -0.615E-04
    4    0.805209E+00     0.100E+01     0.610E-05
    5    0.739921E+00     0.100E+01     0.103E-03
    6   -0.156055E+01     0.100E+01     0.931E-04
    7   -0.704689E+00     0.100E+01    -0.197E-03
    8    0.541099E+00     0.100E+01    -0.277E-03
    9    0.102279E+01     0.100E+01     0.238E-03

   1 EXTRA FUNC. EVALS FOR COVARIANCE AND DIAGNOSTICS.
   1 EXTRA GRAD. EVALS FOR COVARIANCE AND DIAGNOSTICS.

SQRT(RECIPROCAL CONDITION OF (J**T)*RHO"*J) = AT MOST  0.15E-02

COVARIANCE = (J**T * RHO" * J)**-1

ROW  1      0.122E-06
ROW  2      0.637E-05   0.139E-01
ROW  3      0.774E-05   0.743E-02   0.206E-01
ROW  4     -0.484E-05   0.946E-03  -0.166E-02   0.197E-01
ROW  5     -0.615E-05   0.153E-03  -0.299E-02   0.107E-01   0.183E-01
ROW  6      0.189E-04  -0.847E-03  -0.108E-02  -0.475E-02  -0.601E-02
            0.229E-01
ROW  7      0.906E-05   0.105E-02   0.159E-02  -0.912E-03  -0.111E-02
            0.638E-02   0.126E-01
ROW  8     -0.187E-04   0.155E-02  -0.126E-02   0.220E-02   0.379E-02
           -0.303E-02  -0.224E-02   0.156E-01
ROW  9     -0.121E-04  -0.396E-03   0.101E-03   0.321E-02   0.449E-03
           -0.396E-02  -0.127E-02   0.622E-02   0.930E-02

BLOCK REGRESSION DIAGNOSTIC = 0.5 * G(I)**T * H(I)**-1 * H * H(I)**-1 * G(I)
BLOCK  FIRST  LAST    RD(I)         X(I)
    1      1   216    18.3      -0.363146E-02   0.189899      -0.426133
                                 0.824762       0.551742      -0.969219
                                -0.264446       0.333278       1.10424
    2    217   378    13.4      -0.504621E-02  -0.483438E-01  -0.272364
                                 1.00634        0.948275      -2.22987
                                -1.07307        0.739034       1.25643
    3    379   567    7.97      -0.448066E-02  -0.114761E-01  -0.576371
                                 0.687746       0.792413      -1.83698
                                -0.870788       0.577962       0.816515

ASYMPTOTIC T-STATISTICS:
 I               X(I)            T-STAT(I)        STD ERROR
 1  RENT      -0.417176D-02   -0.119269D+02    0.349776D-03
 2  LocD1      0.309478D-02    0.262598D-01    0.117853D+00
 3  LocD2     -0.415382D+00   -0.289368D+01    0.143548D+00
 4  ConD1      0.805209D+00    0.573979D+01    0.140286D+00
 5  ConD2      0.739921D+00    0.546424D+01    0.135412D+00
 6  BedD1     -0.156055D+01   -0.103182D+02    0.151244D+00
 7  BedD2     -0.704689D+00   -0.628199D+01    0.112176D+00
 8  Htype      0.541099D+00    0.433084D+01    0.124941D+00
 9  CDum       0.102279D+01    0.106073D+02    0.964229D-01

NUMBER OF OBSERVATIONS (NOBS) =  567
```

```
LOG-LIKELIHOOD L(EST)  = -0.414996D+03
LOG-LIKELIHOOD L(0)    = -0.622913D+03
-2[L(0) - L(EST)]:     =  0.415834D+03

1 - L(EST)/L(0):       =  0.333782D+00
1 - (L(EST)-NPAR)/L(0) =  0.319334D+00

 (FIXED CHOICE SET SIZE)

AGGREGATE CHOICES AND MARKET SHARES:
  1     121.000  0.2134
  2     133.000  0.2346
  3     313.000  0.5520

STATISTICS FOR CONSTANTS-ONLY MODEL:
   LOG-LIKELIHOOD L(C)    = -0.565715D+03
   -2[L(C) - L(EST)]:     =  0.301438D+03



OUTPUT FOR CONVENIENT RESTART:
RENT
 -0.417176D-02 -0.100000D+03  0.100000D+03
LocD1
  0.309478D-02 -0.100000D+03  0.100000D+03
LocD2
 -0.415382D+00 -0.100000D+03  0.100000D+03
ConD1
  0.805209D+00 -0.100000D+03  0.100000D+03
ConD2
  0.739921D+00 -0.100000D+03  0.100000D+03
BedD1
 -0.156055D+01 -0.100000D+03  0.100000D+03
BedD2
 -0.704689D+00 -0.100000D+03  0.100000D+03
Htype
  0.541099D+00 -0.100000D+03  0.100000D+03
CDum
  0.102279D+01 -0.100000D+03  0.100000D+03
```

(End of output for Example 4.)

## 6. Acknowledgements

This work was supported in part by United States Department of Transportation through the Region Nine Transportation Center at University of California, Berkeley. Dr. David M. Gay and Professor Roy E. Welsch made helpful suggestions for improving this report and the final versions of the programs.

## 7. References

Bunch, D. S. (1991), "Estimability in the Multinomial Probit Model," *Transportation Research B*, Vol 25B(1), February, pp. 1-12.

Bunch, D. S., D. M. Gay, and R. E. Welsch (1991), "Subroutines for Maximum Likelihood and Quasi-likelihood Estimation of Parameters in Nonlinear Regression Models," UC Davis Graduate School of Management Working Paper UCD-GSM-WP0691, September 1991 (Under review at *ACM Transactions of Mathematical Software*).

Bunch, D. S. and R. Kitamura (1990), "Multinomial Probit Estimation Revisited: Testing Estimable Model Specifications, Maximum Likelihood Algorithms, and Probit Integral Approximations for Trinomial Models of Household Car Ownership," Report UCD-TRG-RR-4 (April 1990).

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

Kamakura, W. A. (1989), "The Estimation of Multinomial Probit Models: A New Calibration Algorithm, *Transportation Science*, 23 (4), 253-265.