NLOGIT

Version 5

Reference Guide

by

William H. Greene
Econometric Software, Inc.

N28.12 The Generalized Nested Logit Model

The generalized nested logit model is an extension of the nested logit model in which alternatives may appear in more than one branch. (The behavioral assumptions underlying this model are up to the user.) Alternatives which appear in more than one branch are allocated across branches probabilistically. The model estimated includes the usual nested logit framework (only two levels are supported in this framework), as well as the matrix of allocation parameters. The only difference between this and the more basic nested logit model is the specification of the tree. The model is requested by changing the command name to **GNLOGIT**. Otherwise, the model is the same as the nested logit model. The alternative form

NLOGIT ; GNL ; ...

is also useable. All features of *NLOGIT*, including marginal effects, simulations, etc. are the same as for all other models. The difference here is that when you specify the tree, you may specify that a given alternative appears in more than one branch. (Technical details appear at the end of this section.)

A small example appears below. In this nested logit model, the choice *car* appears in both branches. The probabilities for the allocation are estimated to be .16 and .84. The base case multinomial logit model appears first.

```
GNLOGIT  ; Lhs = mode
         ; Choices = air,train,bus,car
         ; Rhs = one,gc,ttme
         ; Tree = private(air,car), ground(car,train,bus)
         ; Effects: gc(*) $
```

```
--------------------------------------------------------------------------------
Discrete choice (multinomial logit) model
Dependent variable                Choice
Log likelihood function       -199.97662
Estimation based on N =    210, K =   5
Inf.Cr.AIC  =    410.0 AIC/N =    1.952
R2=1-LogL/LogL* Log-L fncn R-sqrd R2Adj
Constants only   -283.7588  .2953 .2862
Chi-squared[ 2]          =    167.56429
Prob [ chi squared > value ] =   .00000
Response data are given as ind. choices
Number of obs.=   210, skipped    0 obs
--------+-----------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
    MODE|  Coefficient       Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
      GC|    -.01578***      .00438    -3.60  .0003     -.02437   -.00719
    TTME|    -.09709***      .01044    -9.30  .0000     -.11754   -.07664
   A_AIR|    5.77636***      .65592     8.81  .0000     4.49078   7.06194
   A_CAR|    3.92300***      .44199     8.88  .0000     3.05671   4.78929
  A_1BUS|    3.21073***      .44965     7.14  .0000     2.32943   4.09204
--------+-----------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
--------------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------------
Generalized Nested Logit Model
Dependent variable                MODE
Log likelihood function      -195.43541
The model has 2 levels.
GNL: Model uses random utility form RU1
Number of obs.=   210, skipped    0 obs
--------+-----------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
    MODE|  Coefficient       Error       z    |z|>Z*         Interval
--------+-----------------------------------------------------------------------
        |Attributes in the Utility Functions (beta)
      GC|    -.02140**       .01030    -2.08  .0379     -.04159   -.00120
    TTME|    -.09368**       .04016    -2.33  .0197     -.17240   -.01496
   A_AIR|    5.30728**      2.67168     1.99  .0470      .07088  10.54367
   A_CAR|    4.21064**      2.00982     2.10  .0362      .27147   8.14980
  A_1BUS|    3.47823**      1.68141     2.07  .0386      .18273   6.77373
        |Dissimilarity parameters. These are mu(branch).
 PRIVATE|    1.95202        1.30315     1.50  .1342     -.60211   4.50615
  GROUND|     .80675         .56368     1.43  .1524     -.29805   1.91155
        |Structural MLOGIT Allocation Model: Constants
tAIR_PRI|        0.0    .....(Fixed Parameter).....
tTRA_GRO|        0.0    .....(Fixed Parameter).....
tBUS_GRO|        0.0    .....(Fixed Parameter).....
tCAR_PRI|   -1.62462      16.42213     -.10  .9212   -33.81141  30.56217
tCAR_GRO|        0.0    .....(Fixed Parameter).....
--------+-----------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
Fixed parameter ... is constrained to equal the value or
had a nonpositive st.error because of an earlier problem.
--------------------------------------------------------------------------------

Generalized Nested Logit
Estimated Allocations of Choices to Branches
Estimated standard errors in parentheses for
allocation values not fixed at 1.0 or 0.0.
        |Branch
--------+-----------------
CHOICE  |PRIVATE  GROUND
--------+--------+--------
AIR      1.0000    .0000
TRAIN     .0000   1.0000
BUS       .0000   1.0000
CAR       .1646    .8354
        ( .0000) ( .0000)
Note: Allocations are multinomial logit
probabilities. Underlying parameters are not
shown in the output:

Elasticity wrt change of X in row choice on Prob[column choice]
--------+-----------------------------------
GC      |     AIR      CAR      CAR    TRAIN
--------+-----------------------------------
     AIR| -1.2007    .6088    .6088    .2953
     CAR|   .6587  -2.5515    .9015    .7905
     CAR|   .3285    .4473  -2.6094    .3941
   TRAIN|   .2449    .7727    .7727  -1.3112
```

Aside from the expanded specification of the tree, the model is otherwise the same as the nested logit model shown earlier. The model contains an allocation matrix,

$$\boldsymbol{\alpha} = [\alpha_{k|j}],$$

which defines the probabilistic allocation of alternatives *k* to branches *j*. The columns of the matrix relate to the branches while the rows refer to the alternatives. The model construction specifies that the rows of the matrix each sum to 1.0. The matrix that was estimated for the model in the example was

```
         |Branch
---------+-----------------
CHOICE   |PRIVATE  GROUND
---------+--------+--------
AIR       1.0000    .0000
TRAIN      .0000   1.0000
BUS        .0000   1.0000
CAR        .1646    .8354
```

The locations of the nonzero entries are specified by the tree definition. In the nested logit model, each row will contain a single 1.0000 and *J*-1 0.0000s. When alternatives appear in more than one branch, then a set of allocation parameters appear in the matrix. These are parameters to be estimated. When there are free parameters to be estimated in $\boldsymbol{\alpha}$, the adding up constraint is imposed by using a multinomial logit form,

$$\alpha_{k|j} = \text{Prob(alternative } k \text{ is in branch } j) = \exp(\theta_{k|j}) / \Sigma_{k,m} \exp(\theta_{j|m}),$$

where the parameters θ are actually estimated by the program. Note the denominator summation is over branches that the alternative appears in. The probabilities sum to one. The identification rule that one of the θs for each alt modeled equals one is imposed. Thus, in the output results above, $\theta_{car,ground} = 0$ and $\theta_{car,private} = -1.625$, so that the probability allocated to the *private* branch is exp(-1.625)/[exp(0)+exp(-1.625)] = 0.1646, which can be seen in the final table of results. You may also specify that these allocations depend on an individual characteristic (not a choice attribute), such as *income*, by using

; GNL = the name of a variable

(Note that even if you use the **GNLOGIT** command, you must have the **; GNL** specification in the command.) In this instance, the multinomial logit probabilities become functions of this variable,

$$\alpha_{k|j} = \text{Prob(alternative } k \text{ is in branch } j) = \exp(\theta_{k|j}+\gamma_{k|j}) / \Sigma_{k,m} \exp(\theta_{j|m}+\gamma_{k|m}).$$

Again, to achieve identification, one of the θs and one of the γs is set equal to zero. The log likelihood function is then assembled from these parameters as follows:

$$Prob(j \mid b) = \frac{\left[\alpha_{j|b} \exp(V_j)\right]^{1/\mu_b}}{\sum_{q=1}^{J} \left[\alpha_{q|b} \exp(V_q)\right]^{1/\mu_b}},$$

$$Prob(b) = \frac{\left\{\sum_{q=1}^{J} \left[\alpha_{q|b} \exp(V_q)\right]^{1/\mu_b}\right\}^{\mu_b}}{\sum_{s=1}^{B} \left\{\sum_{q=1}^{J} \left[\alpha_{q|s} \exp(V_q)\right]^{1/\mu_s}\right\}^{\mu_s}}.$$

Derivatives of this log likelihood function are computed numerically, using two sided finite differences. The BHHH estimator is used for the asymptotic covariance matrix.

N28.13 Box-Cox Nested Logit Model

This variant of the nested logit model allows some attributes to be transformed using the Box-Cox transformation. The model specification adds a degree of flexibility to the functional form. The model specification is the general nested logit form, with

$$U(j) = \sum_{k=1}^{B} \beta_k \left(\frac{x_{jk}^{\lambda_k} - 1}{\lambda_k} \right) + \sum_{m=1}^{K} \beta_m x_{jm} + \sum_{j=1}^{J} \sum_{c=1}^{C} d_{jc} z_c + \varepsilon_j .$$

The utility function contains B attributes, x_{jb} that are transformed, each by an attribute specific transformation parameter, λ_b. It also contains K attributes, x_{jk} that are untransformed – this is the form we have assumed up to this point. Finally, there may be C variables, z_c that are interacted with alternative specific constants. Again, this is the form we have used up to this point. Save for the first term, this is the same model we have used before.

The command setup is

```
NLOGIT      ; Lhs = … ; Choices = …
            ; Tree = specification
            ; Rhs = choice varying attributes
            ; Rh2 = choice invariant characteristics and one
            ; … any other options
            ; Bcl =  list of attributes among the Rhs variables that are
                     subject to the Box-Cox transformation $
```

The utility functions must be in the Rhs/Rh2 format for this specification. An example is

```
NLOGIT      ; Lhs = mode ; Choices = air,train,bus,car
            ; Tree = private(air,car),public(train,bus)
            ; Rhs = gc,invc,invt
            ; Rh2 = one,hinc
            ; Bcl = invc,invt
            ; Effects: gc(*) / invt(*) $
```