USF UNIVERSITY OF SOUTH FLORIDA

Statistical and Econometric Methods

Assignment #5
(Discrete Data – FIML Nested Logit)

As in assignments #3 and #4, you are given 151 observations of a travel survey collected in State College Pennsylvania (same data as in assignment #1). All of the households in the sample are making the morning commute to work. They are all departing from the same origin (a large residential complex in the suburbs) and going to work in the Central Business District. They have the choice of three alternate routes; 1) a four-lane arterial (speed limit = 35mph, 2 lanes each direction), 2) a two-lane rural road (speed limit = 35mph, 1 lane each direction) and 3) a limited access four-lane freeway (speed limit = 55mph, 2 lanes each direction).

Your task, using your best model, is to consider all possible nesting structures and comment on your findings with regard to the value of the logsum and the validity of the nesting structure. The structures to test:

1. The freeway by itself with the arterial and rural roads nested.

2. The arterial by itself with the freeway and rural road nested.

3. The rural road by itself with the freeway and arterial nested.

For reference, see pages 334 to 342 in the text. (Washington, S., M. Karlaftis and F. Mannering (2011) Statistical and econometric methods for transportation data analysis, Second Edition, Chapman & Hall/CRC).

Variables available for your specification are (in file LOGIT-A1.txt):

Variable Number	Explanation
x1	Route chosen, rows: 1 - arterial, 2 - rural road, 3 - freeway
x2	Arterial row indicator; 1 for arterial row, 0 for others
x3	Rural row indicator; 1 for rural row, 0 for others
x4	Freeway row indicator; 1 for freeway row, 0 for others
x5	Traffic flow rate
x6	Number of traffic signals
x7	Distance in tenths of miles
x8	Seat belts: 1 - if wear, 0 - if not
x9	Number of passengers in car
x10	Driver age in years: 1 - 18 to 23, 2 - 24 to 29, 3 - 30 to 39, 4 - 40 to 49, 5 - 50 and above
x11	Gender: 1 - male, 0 - female
x12	Marital status: 1 - single, 0 - married
x13	Number of children
x14	Annual income: 1 - less than 20000, 2 - 20000 to 29999, 3 - 30000 to 39999, 4 - 40000 to 49999, 5 - more than 50000
x15	Model year of car (e.g. 86 = 1986)
x16	Origin of car: 1 - domestic, 0 - foreign
x17	Fuel efficiency in miles per gallon

```
|-> read;nvar=17;nobs=453;file=U:\00Work-Purdue\new_laptop\ LOGIT-A1.txt$
|-> create;cage=86-x15$
|-> nlogit
    ;lhs=x1
    ;choices=arterial,rural,freeway
    ;tree=nf(rural,arterial),three(freeway)
    ;model:
    u(arterial)=dista*x7/
    u(rural)=rural*one+distr*x7/
    u(freeway)=freeway*one+distf*x7/
    u(nf)=malenf*x12+cagenf*cage+domnf*x16
    ;ivset: (three)=[1]$

Iterative procedure has converged
Normal exit:  27 iterations. Status=0, F=    .9479945D+02

------------------------------------------------------------------------------
FIML Nested Multinomial Logit Model
Dependent variable                  X1
Log likelihood function       -94.79945
Restricted log likelihood    -198.93324
Chi squared [  9](P= .000)    208.26758
Significance level              .00000
McFadden Pseudo R-squared      .5234610
Estimation based on N =    151, K =   9
Inf.Cr.AIC  =    207.6 AIC/N =    1.375
--------------------------------------
            Log likelihood R-sqrd R2Adj
No coefficients  -198.9332  .5235 .5088
Constants only   -124.2267  .2369 .2134
At start values  -165.8905  .4285 .4110
Note: R-sqrd = 1 - logL/Logl(constants)
--------------------------------------
Response data are given as ind. choices
The model has 2 levels. Model form: RU1
Coefs. for branch level begin with MALENF
Number of obs.=   151, skipped    0 obs
--------+---------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      X1|  Coefficient       Error       z    |z|>Z*         Interval
--------+---------------------------------------------------------------------
        |Attributes in the Utility Functions (beta).........................
   DISTA|    -.10342***      .03211    -3.22  .0013     -.16637   -.04048
   RURAL|    2.99039**      1.31962     2.27  .0234      .40398   5.57680
   DISTR|    -.15052***      .03624    -4.15  .0000     -.22155   -.07949
 FREEWAY|     .68371        3.71037      .18  .8538    -6.58849   7.95590
   DISTF|    -.16064**       .07611    -2.11  .0348     -.30981   -.01147
        |Attributes of Branch Choice Equations (alpha).....................
  MALENF|     .79933         .68755     1.16  .2450     -.54824   2.14689
  CAGENF|    -.14745**       .06767    -2.18  .0293     -.28008   -.01483
   DOMNF|    -.96785         .73928    -1.31  .1905    -2.41681    .48111
        |IV parameters (RU1), tau(b|l,r),sigma(l|r),phi(r).................
      NF|    1.55789***      .53963     2.89  .0039      .50024   2.61555
   THREE|        1.0    .....(Fixed Parameter).....
--------+---------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
------------------------------------------------------------------------------
```

```
|-> nlogit
    ;lhs=x1
    ;choices=arterial,rural,freeway
    ;tree=na(rural,freeway),three(arterial)
    ;model:
    u(arterial)=artc*one+dista*x7/
    u(rural)=rural*one+distr*x7/
    u(freeway)=distf*x7/
    u(na)=childna*x13+cagena*cage
    ;ivset: (three)=[1]$

Iterative procedure has converged
Normal exit:  23 iterations. Status=0, F=    .9568433D+02

------------------------------------------------------------------------------
FIML Nested Multinomial Logit Model
Dependent variable                  X1
Log likelihood function       -95.68433
Restricted log likelihood    -186.45659
Chi squared [  8](P= .000)    181.54453
Significance level               .00000
McFadden Pseudo R-squared      .4868279
Estimation based on N =    151, K =   8
Inf.Cr.AIC  =    207.4 AIC/N =    1.373
---------------------------------------
            Log likelihood R-sqrd R2Adj
No coefficients  -186.4566  .4868 .4729
Constants only   -124.2267  .2298 .2088
At start values  -165.8905  .4232 .4075
Note: R-sqrd = 1 - logL/Logl(constants)
---------------------------------------
Response data are given as ind. choices
The model has 2 levels. Model form: RU1
Coefs. for branch level begin with CHILDNA
Number of obs.=   151, skipped    0 obs
--------+---------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      X1|  Coefficient       Error       z    |z|>Z*         Interval
--------+---------------------------------------------------------------------
        |Attributes in the Utility Functions (beta).........................
    ARTC|    1.03855        3.58493      .29  .7720    -5.98778   8.06488
   DISTA|    -.11380***      .03180    -3.58  .0003     -.17614   -.05146
   RURAL|    4.17718        3.38511     1.23  .2172    -2.45751  10.81186
   DISTR|    -.18663***      .05522    -3.38  .0007     -.29485   -.07840
   DISTF|    -.10946         .08284    -1.32  .1864     -.27182    .05289
        |Attributes of Branch Choice Equations (alpha).....................
 CHILDNA|    -.25543         .28471     -.90  .3696     -.81344    .30258
  CAGENA|     .14335**       .06754     2.12  .0338      .01098    .27572
        |IV parameters (RU1), tau(b|l,r),sigma(l|r),phi(r)..................
      NA|     .87459**       .38617     2.26  .0235      .11771   1.63146
   THREE|        1.0    .....(Fixed Parameter).....
--------+---------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
------------------------------------------------------------------------------
```

```
|-> nlogit
    ;lhs=x1
    ;choices=arterial,rural,freeway
    ;tree=nr(arterial,freeway),three(rural)
    ;model:
    u(arterial)=artc*one+dista*x7/
    u(rural)=rural*one+distr*x7/
    u(freeway)=distf*x7/
    u(nr)=childnr*x13+cagenr*cage
    ;ivset: (three)=[1]$

Iterative procedure has converged
Normal exit:  28 iterations. Status=0, F=    .9763937D+02

------------------------------------------------------------------------------
FIML Nested Multinomial Logit Model
Dependent variable                    X1
Log likelihood function       -97.63937
Restricted log likelihood    -137.93629
Chi squared [  8](P= .000)     80.59384
Significance level               .00000
McFadden Pseudo R-squared      .2921415
Estimation based on N =    151, K =   8
Inf.Cr.AIC  =    211.3 AIC/N =    1.399
--------------------------------------
            Log likelihood R-sqrd R2Adj
No coefficients  -137.9363  .2921 .2729
Constants only   -124.2267  .2140 .1926
At start values  -165.8905  .4114 .3954
Note: R-sqrd = 1 - logL/Logl(constants)
--------------------------------------
Response data are given as ind. choices
The model has 2 levels. Model form: RU1
Coefs. for branch level begin with CHILDNR
Number of obs.=   151, skipped    0 obs
--------+---------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
      X1|  Coefficient       Error       z    |z|>Z*         Interval
--------+---------------------------------------------------------------------
        |Attributes in the Utility Functions (beta).........................
    ARTC|    2.05344        2.44765      .84  .4015    -2.74387   6.85075
   DISTA|    -.10054***      .02954    -3.40  .0007     -.15844   -.04263
   RURAL|    5.81928*       3.44889     1.69  .0915     -.94042  12.57898
   DISTR|    -.18737***      .03947    -4.75  .0000     -.26473   -.11000
   DISTF|    -.06493         .04442    -1.46  .1438     -.15198    .02213
        |Attributes of Branch Choice Equations (alpha)......................
 CHILDNR|     .32329         .26774     1.21  .2272     -.20147    .84806
  CAGENR|    -.03904         .05186     -.75  .4515     -.14069    .06260
        |IV parameters (RU1), tau(b|l,r),sigma(l|r),phi(r)..................
      NR|    1.39418**        .55494     2.51  .0120      .30651   2.48184
   THREE|        1.0    .....(Fixed Parameter).....
--------+---------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
------------------------------------------------------------------------------
```