LIMDEP TUTORIAL BY MIN AHN

INSTRUCTION FOR ACCESSING AN INSTRUCTOR VOLUME

1. At the ASU PC Network logon you will get a message: "Click OK for the next two requests."

   Click on the **OK** button here.
   Click **OK** and wait 1-2 minutes while the logon scripts execute.
   Click **OK** to get to the sign on screen with the ASU logo displayed over an ASU photograph back ground.

2. At the sign on screen enter your **ASURITE ID** and **password**. Enter both items in lower case.
   Click on **OK**.
   Wait during the message "Mounting AFS volumes." Soon the Window 95 desktop will be displayed.

3. Double click on the **Applications** folder icon on the desktop.

4. Double click on the **Instructor Volumes** folder icon on the desktop.

5. Find the icon named **ECN527**, and double click on it.

6. The U: drive instructor volume is now mounted but you can not see it until the current window is closed. Close the instructor volume window by clicking on **X** in the upper right corner of the window.

7. Double click on the **U: drive** icon on the desktop.

8. Go to the directory Limdep/Program. Click on the **LIMDEP** icon.

9. Now you entered the LIMDEP program (Version 7.0 for windows).

When you have finished using the instructors volume, be sure to LOG OUT so that the next computer user does not have access to your files.

10. Double click on the **Log Out** icon on the desktop. Click on **Log Me Out**. DO NOT turn the computer off.

HOW TO READ DATA

Basic Format:

```
     READ ; NOBS = ...
          ; NVAR = ...
          ; NAMES = ... (THE NAMES OF VARIABLES)
          ; FILE = ... (THE FILE CONTAINING RAW DATA)
          ; FORMAT = ... (SEE LIMDEP MANUAL)      $
```

Example: Using MWPSID82.DB (MW_READ.LIM)

```
     READ ; NOBS=1962; NVAR=25
          ; FILE=MWPSID82.DB
          ; NAMES=  NLF,    EMP,    WRATE,  LRATE,  ED,
                    URB,    MINOR,  AGE,    TENURE, EXP,
                    REGS,   OCCW,   OCCB,   INDUMG, INDUMN,
                    UNION,  UNEMPR, OFINC,  LOFINC, KIDS,
                    WUNE80, HWORK,  USPELL, SEARCH, KIDS5 $

     CREATE; LF = 1 - NLF $
```

(1) This problem is available in MW_READ.LIM. Run the program to read the data,MWPSID82.DB

(2) To save the data, click **File/Project save as**.

PROBIT ESTIMATION

Example: (MW_PROB.LIM)

```
/* Basic Program */

title  ; Employment Probability $
probit ; lhs = emp
       ; rhs = one,ed,urb,minor,lofinc
       ; maxit = 1000
       ; start = 0,0,0,0,0
       ; tlf = 0.00001 ; tlb = 0.00001 ; tlg = 0.00001
       ; alg = newton   ? Can choose bhhh, bfgs, dfp, stedes
       ; margin $       ? estimate p = Pr(y=1) at sample mean

? Saving estimates and covariance matrix
```

```
matrix ; prb = b ; prc = varb $

? Saving the value of likelihood function

calc   ; plogl = logl $

/* Testing hypotheses */

/*
   Wald tests can be computed for any restriction
   LM or LR tests can be computed for certain restrictions:
            equality between parameters or zero restriction

   For example,

        Hypo 1: b(1) = 0, b(2) = b(3)
        Hypo 2: b(1) = 0, b(2) = b(3), b(4) = b(2)+ b(3)
        Hypo 3: b(1) = 0, b(2) = b(3), b(4)^2 = b(5)

  Wald can be used for any of these hypotheses.
  LR or LM can be used only for Hypo 1.

*/

/* Testing hypo 1 */


? Unrestricted Model

namelist ; x   = one,ed,urb,minor,lofinc $
probit   ; lhs = emp
         ; rhs = x
         ; maxit = 1000         $

matrix ; uprb = b ; uprc = varb $
calc   ; ulogl = logl           $

? Restricted Model

probit ; lhs = emp
       ; rhs = x
       ; maxit = 1000
       ; rst = 0,b2,b2,b4,b5 $

matrix ; rprb = b     $
calc   ; rlogl = logl $
```

```
? [Wald Test]

title; Wald test for b1 = 0 and b2 = b3 $
wald ; labels = b1,b2,b3,b4,b5
     ; start = uprb; var = uprc ;
     ; fn1    = b1
     ; fn2    = b2 - b3      $

? [LR test]

title; LR test for b1 = 0 and b2 = b3 $
calc ; list
     ; lrt  = 2*(ulogl - rlogl)
     ; pval = 1 - chi(lrt,2) $

? [LM test]

probit ; lhs   = emp; rhs   = x
       ; start = rprb ; maxit = 0 $

title  ; LM test for b1 = 0 and b2 = b3 $
calc   ; list
       ; lmt = lmstat
       ; pval = 1 - chi(lmt,2)  $

/* Wald for hypo 3 */

title; Wald test for b1 = 0, b2 = b3 and b4^2 = b5  $
wald ; labels = b1,b2,b3,b4,b5
     ; start  = uprb
     ; var    = uprc
     ; fn1    = b1 ;fn2 = b2 - b3; fn3 = b4^2-b5 $

/* Estimating Pr(y=1) and dPr(y=1)/dx_j */

? Estimating Pr(y=1) at mean of x
? Based on unrestricted probit

matrix ; mx  = mean(x) $ ? It is a column vector
calc   ; xb1 = mx(1,1) ; xb2 = mx(2,1); xb3 = mx(3,1)
       ; xb4 = mx(4,1) ; xb5 = mx(5,1)   $

title  ; Pr(emp=1) at means of regressors $
wald   ; labels = b1,b2,b3,b4,b5; start = uprb; var = uprc;
       ; fn1    = phi(xb1*b1+xb2*b2+xb3*b3+xb4*b4+xb5*b5) $
```

```
? Estimating dPr(y=1)/dx

title  ; Marginal effects  $
wald   ; labels = b1,b2,b3,b4,b5
       ; start  = uprb
       ; var    = uprc
       ; fn1    = xb1*b1+xb2*b2+xb3*b3+xb4*b4+xb5*b5
       ; fn2    = n01(fn1)*b1
       ; fn3    = n01(fn1)*b2
       ; fn4    = n01(fn1)*b3   $
```

Output:

```
--> RESET
--> LOAD;file="C:\Aaschool\ECON727\DATA\MW.lpj"$
LOAD has reconstructed your previous session.
--> title  ; Employment Probability
    probit ; lhs = emp
    ; rhs = one,ed,urb,minor,lofinc
    ; maxit = 1000
    ; start = 0,0,0,0,0
    ; tlf = 0.00001 ; tlb = 0.00001 ; tbg = 0.00001
    ; alg = newton   ? Can choose bhhh, bfgs, dfp, stedes
    ; margin $       ? estimate p = Pr(y=1) at sample mean
--> matrix ; prb = b ; prc = varb $
--> calc   ; plogl = logl $
--> namelist ; x   = one,ed,urb,minor,lofinc $
--> probit   ; lhs = emp
    ; rhs = x
    ; maxit = 1000        $


+-----------------------------------------------------------------------+
| Dependent variable is binary, y=0 or y not equal 0                    |
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = EMP      Mean=   .4704383282    , S.D.=   .4992525893     |
| Model size: Observations =    1962, Parameters =   5, Deg.Fr.=   1957 |
| Residuals:  Sum of squares= 474.9962035    , Std.Dev.=         .49266 |
| Fit:        R-squared=  .028211, Adjusted R-squared =          .02622 |
| Model test: F[  4,   1957] =   14.20,    Prob value =          .00000 |
| Diagnostic: Log-L =  -1392.4945, Restricted(b=0) Log-L =   -1420.5675 |
|             LogAmemiyaPrCrt.=   -1.413, Akaike Info. Crt.=      1.425 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  .9170578290      .17362375        5.282   .0000
 ED        .2915849316E-01  .51513883E-02    5.660   .0000  12.205403
 URB       .6459244774E-01  .24441441E-01    2.643   .0082  .68654434
 MINOR     .2566128748E-01  .26742220E-01     .960   .3373  .27166157
 LOFINC   -.8619967756E-01  .17367289E-01   -4.963   .0000  9.9052275

Normal exit from iterations. Exit status=0.
```

```
               +--------------------------------------------+
               | Binomial Probit Model                      |
               | Maximum Likelihood Estimates               |
               | Dependent variable                 EMP     |
               | Weighting variable                 ONE     |
               | Number of observations            1962     |
               | Iterations completed                 4     |
               | Log likelihood function      -1327.781     |
               | Restricted log likelihood    -1356.524     |
               | Chi-squared                   57.48581     |
               | Degrees of freedom                   4     |
               | Significance level            .0000000     |
               +--------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Index function for probability
 Constant  1.228085188      .47320614        2.595   .0095
 ED        .7629799963E-01  .13446644E-01    5.674   .0000  12.205403
 URB       .1708719958      .63111585E-01    2.707   .0068  .68654434
 MINOR     .5968648873E-01  .68899573E-01     .866   .3863  .27166157
 LOFINC   -.2390755461      .48122027E-01   -4.968   .0000  9.9052275


Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.

            Predicted
------  ----------  +  -----
Actual      0    1  |  Total
------  ----------  +  -----
  0       740  299  |   1039
  1       527  396  |    923
------  ----------  +  -----
Total    1267  695  |   1962
```

--> matrix ; uprb = b ; uprc = varb $
--> calc   ; ulogl = logl $
--> probit ; lhs = emp
**   ; rhs = x**
**   ; maxit = 1000**
**   ; rst = 0,b2,b2,b4,b5 $**

```
+-----------------------------------------------------------------------+
| Dependent variable is binary, y=0 or y not equal 0                    |
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = EMP      Mean=   .4704383282    , S.D.=   .4992525893     |
| Model size: Observations =    1962, Parameters =   5, Deg.Fr.=   1957 |
| Residuals:  Sum of squares= 474.9962035    , Std.Dev.=         .49266 |
| Fit:        R-squared=  .028211, Adjusted R-squared =          .02622 |
| Model test: F[  4,    1957] =   14.20,    Prob value =         .00000 |
| Diagnostic: Log-L =  -1392.4945, Restricted(b=0) Log-L =   -1420.5675 |
|             LogAmemiyaPrCrt.=   -1.413, Akaike Info. Crt.=      1.425 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  .9170578290      .17362375        5.282   .0000
 ED        .2915849316E-01  .51513883E-02    5.660   .0000  12.205403
 URB       .6459244774E-01  .24441441E-01    2.643   .0082  .68654434
 MINOR     .2566128748E-01  .26742220E-01     .960   .3373  .27166157
 LOFINC   -.8619967756E-01  .17367289E-01   -4.963   .0000  9.9052275

Normal exit from iterations. Exit status=0.
```

```
              +---------------------------------------------+
              | Binomial Probit Model                       |
              | Maximum Likelihood Estimates                |
              | Dependent variable                  EMP     |
              | Weighting variable                  ONE     |
              | Number of observations             1962     |
              | Iterations completed                  4     |
              | Log likelihood function       -1331.862     |
              | Restricted log likelihood     -1356.524     |
              | Chi-squared                    49.32278     |
              | Degrees of freedom                    2     |
              | Significance level             .0000000     |
              +---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Index function for probability
 Constant  .0000000000    ........(Fixed Parameter)........
 ED        .8579347233E-01  .12778130E-01    6.714   .0000  12.205403
 URB       .8579347233E-01  .12778130E-01    6.714   .0000  .68654434
 MINOR     .1336072788      .63756953E-01    2.096   .0361  .27166157
 LOFINC   -.1233762049      .17163930E-01   -7.188   .0000  9.9052275


Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.

            Predicted
------  ----------  +  -----
Actual      0    1  |  Total
------  ----------  +  -----
  0       756  283  |   1039
  1       571  352  |    923
------  ----------  +  -----
Total    1327  635  |   1962
```

```
--> matrix ; rprb = b     $
--> calc   ; rlogl = logl $
--> title; Wald test for b1 = 0 and b2 = b3 $
--> wald ; labels = b1,b2,b3,b4,b5
    ; start  = uprb
    ; var    = uprc
    ; fn1    = b1
    ; fn2    = b2 - b3      $
```

```
               +-----------------------------------------------+
               | WALD procedure. Estimates and standard errors |
               | for nonlinear functions and joint test of     |
               | nonlinear restrictions.                       |
               | Wald Statistic            =      8.06527      |
               | Prob. from Chi-squared[ 2] =      .01773      |
               +-----------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Fncn( 1)  1.228085188      .47320614        2.595   .0095
 Fncn( 2) -.9457399616E-01  .65465798E-01   -1.445   .1486
```

```
--> title; LR test for b1 = 0 and b2 = b3 $
--> calc ; list
    ; lrt  = 2*(ulogl - rlogl)
    ; pval = 1 - chi(lrt,2) $
    LRT     =  .81630295544032380D+01
    PVAL    =  .16881874007718900D-01
```

```
--> probit ; lhs   = emp
    ; rhs   = x
    ; start = rprb
    ; maxit = 0 $

+-----------------------------------------------------------------------+
| Dependent variable is binary, y=0 or y not equal 0                    |
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = EMP      Mean=   .4704383282    , S.D.=   .4992525893     |
| Model size: Observations =    1962, Parameters =   5, Deg.Fr.=   1957 |
| Residuals:  Sum of squares= 474.9962035    , Std.Dev.=         .49266 |
| Fit:        R-squared=  .028211, Adjusted R-squared =          .02622 |
| Model test: F[  4,   1957] =   14.20,    Prob value =          .00000 |
| Diagnostic: Log-L =  -1392.4945, Restricted(b=0) Log-L =   -1420.5675 |
|             LogAmemiyaPrCrt.=   -1.413, Akaike Info. Crt.=      1.425 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  .9170578290      .17362375        5.282   .0000
 ED        .2915849316E-01  .51513883E-02    5.660   .0000  12.205403
 URB       .6459244774E-01  .24441441E-01    2.643   .0082  .68654434
 MINOR     .2566128748E-01  .26742220E-01     .960   .3373  .27166157
 LOFINC   -.8619967756E-01  .17367289E-01   -4.963   .0000  9.9052275

Maximum iterations reached. Exit iterations with status=1.
Maxit = 0. Computing LM statistic at starting values.
No iterations computed and no parameter update done.
              +---------------------------------------------+
              | Binomial Probit Model                       |
              | Maximum Likelihood Estimates                |
              | Dependent variable                 EMP      |
              | Weighting variable                 ONE      |
              | Number of observations            1962      |
              | Iterations completed                 1      |
              | LM Stat. at start values      8.027283      |
              | LM statistic kept as scalar    LMSTAT       |
              | Log likelihood function      -1331.862      |
              | Restricted log likelihood    -1356.524      |
              | Chi-squared                   49.32278      |
              | Degrees of freedom                   4      |
              | Significance level            .0000000      |
              +---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Index function for probability
 Constant  .0000000000      .46145275         .000  1.0000
 ED        .8579347233E-01  .13444274E-01    6.381   .0000  12.205403
 URB       .8579347233E-01  .62895744E-01    1.364   .1725  .68654434
 MINOR     .1336072788      .68714027E-01    1.944   .0518  .27166157
 LOFINC   -.1233762049      .46617697E-01   -2.647   .0081  9.9052275


Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.

            Predicted
------  ----------  +  -----
Actual      0    1  |  Total
------  ----------  +  -----
  0       756  283  |   1039
  1       571  352  |    923
------  ----------  +  -----
Total    1327  635  |   1962
```

```
--> title  ; LM test for b1 = 0 and b2 = b3 $
--> calc   ; list
    ; lmt = lmstat
    ; pval = 1 - chi(lmt,2)  $

    LMT      =  .80272833001843540D+01
    PVAL     =  .18067479861156980D-01


--> title; Wald test for b1 = 0, b2 = b3 and b4^4 = b5  $
--> wald ; labels = b1,b2,b3,b4,b5
    ; start  = uprb
    ; var    = uprc
    ; fn1    = b1
    ; fn2    = b2 - b3
    ; fn3    = b4^2 - b5  $

               +-----------------------------------------------+
               | WALD procedure. Estimates and standard errors |
               | for nonlinear functions and joint test of     |
               | nonlinear restrictions.                        |
               | Wald Statistic             =     52.38105     |
               | Prob. from Chi-squared[ 3] =       .00000     |
               +-----------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Fncn( 1)  1.228085188      .47320614        2.595   .0095
 Fncn( 2) -.9457399616E-01  .65465798E-01   -1.445   .1486
 Fncn( 3)  .2426380230      .46728575E-01    5.192   .0000


--> matrix ; mx  = mean(x) $ ? It is a column vector
--> calc   ; xb1 = mx(1,1)
    ; xb2 = mx(2,1)
    ; xb3 = mx(3,1)
    ; xb4 = mx(4,1)
    ; xb5 = mx(5,1)    $
--> title  ; Pr(emp=1) at means of regressors $
--> wald   ; labels = b1,b2,b3,b4,b5
    ; start  = uprb
    ; var    = uprc
    ; fn1    = phi(xb1*b1+xb2*b2+xb3*b3+xb4*b4+xb5*b5) $


               +-----------------------------------------------+
               | WALD procedure. Estimates and standard errors |
               | for nonlinear functions and joint test of     |
               | nonlinear restrictions.                        |
               | Wald Statistic             =   1710.21762     |
               | Prob. from Chi-squared[ 1] =       .00000     |
               +-----------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Fncn( 1)  .4700122978      .11365369E-01   41.355   .0000


--> title  ; Marginal effects  $
--> wald   ; labels = b1,b2,b3,b4,b5
    ; start  = uprb
    ; var    = uprc
    ; fn1    = xb1*b1+xb2*b2+xb3*b3+xb4*b4+xb5*b5
    ; fn2    = n01(fn1)*b1
    ; fn3    = n01(fn1)*b2
    ; fn4    = n01(fn1)*b3    $
```

```
+--------------------------------------------------+
| WALD procedure. Estimates and standard errors    |
| for nonlinear functions and joint test of        |
| nonlinear restrictions.                          |
| Wald Statistic              =     60.67600       |
| Prob. from Chi-squared[ 4] =       .00000       |
+--------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Fncn( 1) -.7523894869E-01  .28569505E-01   -2.634   .0084
 Fncn( 2)  .4885503289      .18831825        2.594   .0095
 Fncn( 3)  .3035246511E-01  .53493381E-02    5.674   .0000
 Fncn( 4)  .6797539012E-01  .25106007E-01    2.708   .0068
```

[EXERCISE]Use MWPSID82.DB. Do probit for:

LHS=EMP,
RHS=ONE, ED, URB, MINOR, AGE, REGS, UNEMPR, LOFINC, KIDS5, EXP.

Construct Wald, LR and LM statistics for H_o: no effect of other family income, and the effect of AGE = the effect of EXP.

[EXERCISE]Use MWPSID82.DB. Do PROBIT as you did above. Estimate the same model by BHHH. Compare your new result with the result from the above program.

INFORMATION ON MWPSID82.DB

This is the data set of married women in 1981 sampled from PSID. Total number of observations are 1962, and 25 variables are observed.

VARIABLES	DEFINITION
NLF	NLF=1 IF NON-LABOR-FORCE (HOUSEWIFE)
EMP	EMP=1 IF EMPLOYED
WRATE	HOURLY WAGE RATE ($)
LRATE	Log of WRATE = LOG(WRATE+1)
ED	YEARS OF EDUCATION
URB	URB=1 IF RESIDENT IN SMSA
MINOR	MINOR=1 IF BLACK AND HISPANIC
AGE	YEARS OF AGE
TENURE	MONTHS UNDER THE CURRENT EMPLOYER
EXP	NUMBER OF YEARS WORKED SINCE AGE 18
REGS	REGS=1 IF LIVES IN THE SOUTH OF U.S.
OCCW	OCCW=1 IF WHITE COLOR
OCCB	OCCB=1 IF BLUE COLOR
INDUMG	INDUMG=1 IF IN THE MANUFACTURING INDUSTRY
INDUMN	INDUMN=1 IF NOT IN MANUFACTURING SECTOR
UNION	UNION=1 IF UNION MEMBER
UNEMPR	% UNEMPLOYMENT RATE IN THE RESIDENT'S COUNTY, 1980
OFINC	OTHER FAMILY MEMBER'S INCOME IN 1980 ($)
LOFINC	LOG OF (OFINC+1)
KIDS	NUMBER OF CHILDREN 17 YEARS OF AGE
HWORK	HOURS OF HOMEWORK PER WEEK
USPELL	UNEMPLOYED WEEKS FOR EMPLOYED WIFE
SEARCH	WEEKS LOOKING FOR JOB IN 1980
WUNE80	ACTUAL UNEMPLOYED HOURS
KIDS5	NUMBER OF CHILDREN 5 YEARS OF AGE

MAXIMUM SCORE ESTIMATION

MSE is designed for relatively small problems. The initial limits are 15 parameters and 2,000 observations.

Example 1: (MSE1.LIM)

namelist ; x = one,ed,urb,minor,lofinc $

MSCORE ; LHS = emp; RHS = X; NBT = 100; MAXIT=1000 $

Ouput

```
+--------------------------------------------------------------------+
| Maximum Score Estimates of Linear Quantile                         |
| Regression Model from Binary Response Data                         |
| Quantile                .500      Number of Parameters =     5     |
| Observations input   =  1962      Maximum Iterations   =  1000     |
| End Game Iterations  =   100      Bootstrap Estimates  =    20     |
| Check Ties?              No                                        |
| Save bootstraps?         No                                        |
| Start values from MSCORE (normalized)                              |
| Normal exit after  100 iterations.                                 |
| Score functions:     Naive   At theta(0)      Maximum             |
|            Raw      .05912       .05912         .17533             |
|    Normalized      .52956       .52956         .58767             |
| Estimated MSEs from  20 bootstrap samples                          |
| (Nonconvergence in   0 cases)                                      |
| Angular deviation (radians) of bootstraps from estimate            |
| Mean square =   .881535             Mean absolute =   .760578      |
| Standard errors below are based on bootstrap mean squared          |
| deviations.  These and the t-ratios are only approximations.       |
+--------------------------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Constant	-.7094313341	.80360150	-.883	.3773	
ED	.3659168155	.12119087	3.019	.0025	12.205403
URB	.4329113018	.15678261	2.761	.0058	.68654434
MINOR	-.1325863919	.29330609	-.452	.6512	.27166157
LOFINC	-.3972665601	.16375789	-2.426	.0153	9.9052275

```
Frequencies of actual & predicted outcomes
Predicted outcome is the sign of x(i)*theta.

            Predicted
------  ----------  +  -----
Actual       0    1 |  Total
------  ----------  +  -----
  0        558  481 |   1039
  1        328  595 |    923
------  ----------  +  -----
Total      886 1076 |   1962
```

Example 2: (MSE2.LIM)

```
/*

  This program is for Monte Carlo Simulation:
        Model: ys = b1 + b2*x2 + e, b1 = b2 = 1/SQR(2)

  Compare the finite-sample biases
         in estimates b2 by MSE(MSCORE) and Probit.

*/

? Define Sample Size

CALC   ; SS = 100 $
SAMPLE ; 1 - SS   $

? Define # of simulations

CALC   ; SN = 10  $

? VALUE FOR B2

CALC  ; TB2 = 1/SQR(2) $

? INITIAL VALUE FOR BIAS BY MSE, PROBIT AND LOGIT
CALC ; BIASMB2 = 0; BIASPB2 = 0   $
```

```
? BEGIN THE PROCEDURE

PROCEDURE

CREATE    ; X2 = RNN(1,1) ; E = RNN(0,1) $
CREATE    ; YS = (1/SQR(2)) + (1/SQR(2))*X2 + E $
CREATE    ; Y = YS>0        $

NAMELIST  ; X = ONE,X2 $

? MSE
MSCORE  ; LHS=Y ; RHS=X $
CALC    ; BIASMB2 = BIASMB2 + (B(2,1)-TB2) $

? PROBIT

PROBIT  ; LHS=Y ; RHS=X $
CALC    ; BIASPB2 = BIASPB2 + (B(2,1)-TB2) $

ENDPROCEDURE

EXECUTE ; SILENT ; I = 1, SN $

CALC ; LIST
     ; BIASM = BIASMB2/SN
     ; BIASP = BIASPB2/SN  $
```

[EXERCISE] Run the program with SN = 100. Which estimator has smaller bias? Changing SS = 50, 100 and 1000, see how the biases of MSE and Probit estimates change.

ORDERED PROBIT

Program for reading ORD_DATA.DB: (ORD_READ.LIM)

```
READ ; NVAR=3 ; NOBS=500
     ; FILE = ORD_DATA.DB; FORMAT=(6G13.6)
     ; NAMES = X1,X2,Y $
```

Program for Ordered Probit: (ORD_PROB.LIM)

```
NAMELIST  ; X = ONE,X1,X2 $
ORDERED   ; LHS=Y ; RHS=X ; MARGIN ; PAR $
```

Output:

```
              +---------------------------------------------+
              | Ordered Probit Model                        |
              | Maximum Likelihood Estimates                |
              | Dependent variable                 Y        |
              | Weighting variable               ONE        |
              | Number of observations           500        |
              | Iterations completed              13        |
              | Log likelihood function    -62.40438        |
              | Restricted log likelihood  -512.2856        |
              | Chi-squared                 899.7625        |
              | Degrees of freedom                 2        |
              | Significance level          .0000000        |
              |    Cell frequencies for outcomes            |
              |  Y Count Freq  Y Count Freq  Y Count Freq   |
              |  0   197 .394  1    35 .070  2    19 .038   |
              |  3   249 .498                               |
              +---------------------------------------------+
+---------+-------------+---------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+---------------+--------+---------+----------+
          Index function for probability
 Constant  1.910662552      .41836551        4.567   .0000
 X1        2.070909800      .32510834        6.370   .0000  .91768402
 X2        .5310649242      .96543109E-01    5.501   .0000  .76713340
          Threshold parameters for index
 Mu( 1)    3.319234065      .66037247        5.026   .0000
 Mu( 2)    4.917539009      .90157315        5.454   .0000


            +-------------------------------------------------------+
            | Marginal Effects for OrdProbt                         |
            +----------+----------+----------+----------+----------+
            | Variable | Y=0      | Y=1      | Y=2      | Y=3      |
            +----------+----------+----------+----------+----------+
            | ONE      |   -.0001 |   -.5086 |   -.0883 |    .5970 |
            | X1       |   -.0001 |   -.5513 |   -.0957 |    .6471 |
            | X2       |    .0000 |   -.1414 |   -.0245 |    .1659 |
            +----------+----------+----------+----------+----------+

Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.
```

```
             Predicted
------  --------------------  +  -----
Actual      0    1    2    3  |  Total
------  --------------------  +  -----
  0       194    3    0    0  |    197
  1         4   25    6    0  |     35
  2         0    5    8    6  |     19
  3         0    0    4  245  |    249
------  --------------------  +  -----
Total     198   33   18  251  |    500
```

Program for Ordered Probit with HET: (ORD_PROB.LIM)

```
NAMELIST   ; X = ONE,X1,X2 $
ORDERED    ; LHS=Y ; RHS=X ; HET ; RH2 = X1,x2 ; MARGIN ; PAR $
```

Output:

```
                +---------------------------------------------+
                | Ordered Probit Model                        |
                | Maximum Likelihood Estimates                |
                | Dependent variable                   Y      |
                | Weighting variable                 ONE      |
                | Number of observations             500      |
                | Iterations completed                21      |
                | Log likelihood function      -61.69838      |
                | Restricted log likelihood    -512.2856      |
                | Chi-squared                   901.1745      |
                | Degrees of freedom                   4      |
                | Significance level            .0000000      |
                |    Cell frequencies for outcomes            |
                |  Y Count Freq  Y Count Freq  Y Count Freq   |
                |  0   197 .394  1    35 .070  2    19 .038   |
                |  3   249 .498                               |
                | Terms   4 to   5 are for variance.          |
                +---------------------------------------------+
+---------+-------------+---------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+---------------+--------+---------+----------+
          Index function for probability
 Constant  1.952765793      .50647697       3.856   .0001
 X1        2.202009071      .42735794       5.153   .0000   .91768402
 X2         .5724646749     .13028472       4.394   .0000   .76713340
          Variance function
 X1       -.3920695036E-01  .15416593       -.254   .7993   .91768402
 X2         .2509306915E-01  .47567677E-01    .528   .5978   .76713340
          Threshold parameters for index
 Mu( 1)    3.507495686      .81328128       4.313   .0000
 Mu( 2)    5.161587990      1.1134873       4.636   .0000
```

```
              +-------------------------------------------------------+
              | Marginal Effects for OrdProbt                         |
              +----------+----------+----------+----------+----------+
              | Variable | Y=0      | Y=1      | Y=2      | Y=3      |
              +----------+----------+----------+----------+----------+
              | ONE      |    .0000 |   -.5186 |   -.0741 |    .5928 |
              | X1       |    .0000 |   -.5848 |   -.0836 |    .6684 |
              | X2       |    .0000 |   -.1520 |   -.0217 |    .1738 |
              | X1       |    .0001 |    .4693 |   -.9134 |    .4439 |
              | X2       |    .0002 |    .5292 |  -1.0300 |    .5006 |
              +----------+----------+----------+----------+----------+
```

```
Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.

             Predicted
------  --------------------  +  -----
Actual      0   __1    2    3  |  Total
------  --------------------  +  -----
  0       194    3    0    0  |    197
  1         4   25    6    0  |     35
  2         0    5    8    6  |     19
  3         0    0    4  245  |    249
------  --------------------  +  -----
Total     198   33   18  251  |    500
```

[EXERCISE]

1) Use ORD_DATA.DB. Test for the existence of HET by Wald, LR and LM.
2) Make a program which computes the marginal effects of X1 on Pr(y=1) at the sample means of X1 and X2.

MULTINOMIAL LOGIT

Example: (MW_MLM.LIM)

```
? Creating variable for occupation
? OCC = 0, if service; = 1, if white; = 2, if blue.

CREATE ; OCC = OCCW + 2*OCCB $

? Choose employed only

REJECT; EMP = 0 $

? Do MLM
? Coefficients for OCC = 0 are set to zeros.

NAMELIST; X = ONE,ED,URB,MINOR,LOFINC $
LOGIT ; LHS = OCC ; RHS = X ; MARGIN  $

? MLM with restriction b_1 = b_2

CALC  ; K = COL(X) $
LOGIT ; LHS = OCC ; RHS = X
      ; RST = K_B,K_B $

SAMPLE ; ALL $
```

Output:

```
               +---------------------------------------------+
               | Multinomial Logit Model                     |
               | Maximum Likelihood Estimates                |
               | Dependent variable                 OCC      |
               | Weighting variable                 ONE      |
               | Number of observations             923      |
               | Iterations completed                 7      |
               | Log likelihood function      -671.8235      |
               | Restricted log likelihood    -841.1677      |
               | Chi-squared                   338.6885      |
               | Degrees of freedom                   8      |
               | Significance level            .0000000      |
               +---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Characteristics in numerator of Prob[Y = 1]
 Constant -7.718311837      1.7357637       -4.447   .0000
 ED        .6675239548      .70670487E-01    9.446   .0000  12.457205
 URB       .2777434765      .21712192        1.279   .2008  .71614301
 MINOR    -.9495632405      .20731454       -4.580   .0000  .28494041
 LOFINC    .7928185795E-01  .15593235         .508   .6111  9.8432227
          Characteristics in numerator of Prob[Y = 2]
 Constant  4.661300540      1.7791433        2.620   .0088
 ED       -.1317677880      .58362277E-01   -2.258   .0240  12.457205
 URB      -.5636828799      .23791127       -2.369   .0178  .71614301
 MINOR    -.4166512440      .23969726       -1.738   .0822  .28494041
 LOFINC   -.2989224821      .17310528       -1.727   .0842  9.8432227


+-------------------------------------------+
| Partial derivatives of probabilities with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs.  |
| A full set is given for the entire set of |
| outcomes, OCC      = 0 to OCC      =  2.  |
| Probabilities at the mean vector are      |
|  0= .188 1= .692 2= .120                  |
+-------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Marginal effects on Prob[Y = 0]
 Constant  .8970085202      .25111906        3.572   .0004
 ED       -.8369597078E-01  .95455832E-02   -8.768   .0000  12.457205
 URB      -.2335294950E-01  .31386949E-01    -.744   .4569  .71614301
 MINOR     .1326770374      .29953771E-01    4.429   .0000  .28494041
 LOFINC   -.3554697100E-02  .22870034E-01    -.155   .8765  9.8432227
          Marginal effects on Prob[Y = 1]
 Constant -2.032158327      .31222880       -6.509   .0000
 ED        .1531730201      .13409235E-01   11.423   .0000  12.457205
 URB       .1060726545      .39922710E-01    2.657   .0079  .71614301
 MINOR    -.1676263157      .39497910E-01   -4.244   .0000  .28494041
 LOFINC    .4176258361E-01  .26774137E-01    1.560   .1188  9.8432227
          Marginal effects on Prob[Y = 2]
 Constant  1.135149806      .16732040        6.784   .0000
 ED       -.6947704936E-01  .71452164E-02   -9.724   .0000  12.457205
 URB      -.8271970500E-01  .22400671E-01   -3.693   .0002  .71614301
 MINOR     .3494927833E-01  .23443352E-01    1.491   .1360  .28494041
 LOFINC   -.3820788651E-01  .14389635E-01   -2.655   .0079  9.8432227

Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.
```

Actual	Predicted 0	1	2	Total
0	33	131	24	188
1	14	557	12	583
2	27	85	40	152
Total	74	773	76	923

Normal exit from iterations. Exit status=0.

```
+--------------------------------------------+
| Multinomial Logit Model                    |
| Maximum Likelihood Estimates               |
| Dependent variable                   OCC   |
| Weighting variable                   ONE   |
| Number of observations               923   |
| Iterations completed                   5   |
| Log likelihood function       -930.7384   |
+--------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Characteristics in numerator of Prob[Y = 1]					
Constant	-2.246577102	1.4168164	-1.586	.1128	
ED	.2945460414	.44213261E-01	6.662	.0000	12.457205
URB	-.6619168682E-01	.19853983	-.333	.7388	.71614301
MINOR	-.7559059112	.18812078	-4.018	.0001	.28494041
LOFINC	-.3209143823E-01	.14166104	-.227	.8208	9.8432227
Characteristics in numerator of Prob[Y = 2]					
Constant	-2.246577102	1.4168164	-1.586	.1128	
ED	.2945460414	.44213261E-01	6.662	.0000	12.457205
URB	-.6619168682E-01	.19853983	-.333	.7388	.71614301
MINOR	-.7559059112	.18812078	-4.018	.0001	.28494041
LOFINC	-.3209143823E-01	.14166104	-.227	.8208	9.8432227

Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.

Actual	Predicted 0	1	2	Total
0	50	138	0	188
1	16	567	0	583
2	48	104	0	152
Total	114	809	0	923

[Exercise] Use MWPSID82.DB. Do MLM for:

LHS = OCC
RHS = ONE, ED,URB,MINOR,LOFINC,AGE,EXP,KIDS5.

1) Test H_o: Each women is equally likely to be a white- or blue-collar worker. Use LM, LR and Wald.
2) Estimate the probability of being a white collar at sample means of regressors.

BIVARIATE PROBIT MODEL

Consider the following model: (same as in the class)

$Y1^{*} = {}_{11} + {}_{12}X12 + {}_{13}X13$
$Y2^{*} = {}_{21} + {}_{22}X22 + {}_{23}X23$

$Y1 = 1$ IF $Y1^{*} >= 0$; $Y1 = 0$, OTHERWISE
$Y2 = 1$ IF $Y2^{*} >= 0$; $Y2 = 0$, OTHERWISE

Full Observability:

```
NAMELIST ; X1 = ONE,X12,X13
         ; X2 = ONE,X22,X23 $

? CREATE INITIAL VALUES FOR BIVARIATE PROBIT

PROBIT   ; LHS = Y1 ; RHS = X1  $
MATRIX   ; BT  = B              $
PROBIT   ; LHS = Y2 ; RHS = X2  $
MATRIX   ; BP  = B              $

? MLE FOR FULL OBSERVABILITY MODEL

BIVA ; LHS = Y1,Y2 ; RH1 = X1 ; RH2 = X2
     ; START = BT,BP,0
     ; TLF = 0.00001 ; TLB = 0.00001 ; TLG = 0.00001
     ; MAXIT = 1000
     ; PAR $
```

COMMENT: The first part of the program creates initial values for the bivariate Probit. In BIVA command, you have to specify the initial values using Subcommand, "START = ...". If you omit "PAR", LIMDEP store the estimates of 's and Separately. is stored as a scaler in the name of "RHO". This is not good when you want to test some hypotheses using all estimates. If you use "PAR", all estimates of are are stored in the matrix, B, and covariance matrix of all the estimated parameters in VARB. It is better idea to use "PAR". Whether you use "PAR" or not, is stored as a scaler.

Censored Probit

Y2 is observed only if Y1 = 1.

```
NAMELIST ; X1 = ONE,X12,X13
         ; X2 = ONE,X22,X23 $
```

```
? CREATE INITIAL VALUES FOR BIVARIATE PROBIT

PROBIT    ; LHS = Y1 ; RHS = X1 $
MATRIX    ; BT  = B $
REJECT    ; Y1 # 1 $
PROBIT    ; LHS = Y2 ; RHS = X2  $
MATRIX    ; BP  = B $
SAMPLE    ; ALL $

? CENSORED PROBIT MODEL

BIVA      ; LHS = Y2,Y1 ; RH1 = X2 ; RH2 = X1
          ; START = BP,BT,0
          ; SELECTION
          ; TLF = 0.00001 ; TLB = 0.00001 ; TLG = 0.00001
          ; MAXIT = 1000
          ; PAR $
```

COMMENT: The initial values for 's are created by separate Probits. These Probit estimates are efficient if is equal to zero. For the censored Probit, You just include "SELECTION" in the bivariate Probit command. You also note that LIMDEP reports the estimates in order of $_2$, $_1$ and . In the case of full observability, Poirier, Abowd-Farber, LIMDEP Reports $_1$, $_2$ and .

Poirier probit

Only Y = Y1*Y2 are observed.

```
NAMELIST ; X1 = ONE,X12,X13
         ; X2 = ONE,X22,X23 $
PROBIT   ; LHS = Y ; RH1 = X1 ; RH2 = X2
         ; TLF = 0.00001 ; TLB = 0.00001 ; TLG = 0.00001
         ; MAXIT = 1000
         ; PAR $
```

Comment: You don't have to give initial values.

Abowed-Farber Probit

```
NAMELIST ; X1 = ONE,X12,X13
         ; X2 = ONE,X22,X23 $
PROBIT   ; LHS = Y ; RH1 = X1 ; RH2 = X2 ; SELECTION
         ; TLF = 0.00001 ; TLB = 0.00001 ; TLG = 0.00001
         ; MAXIT = 1000 $
```

SURE (Seemingly Uncorrelated Regression Model)

Suppose you observe $Y1^*$(=YY1) AND $Y2^*$ (=YY2). Then, you can estimate 's by SURE method. FOR this, use following commands:

```
NAMELIST ; X1 = ONE,X12,X13
         ; X2 = ONE,X22,X23 $
SURE     ; LHS = YY1,YY2
         ; EQ1 = X1
         ; EQ2 = X2 $
```

Information on BIV.DB

Data in BIV.DB is generated by following:

```
SAMPLE ; 1-500 $
CALC   ; SS=SQR(3)/2 $
CREATE ; V1=RNN(0,1)
       ; V2=RNN(0,SS)
       ; V3=RNN(0,2)
       ; E1=V1 ; E2=V1/2+V2 ; E3=V3+V1
       ; X1=RNN(-1,2) ; X2=RNN(0,2) ; X3=RNN(3,2)
       ; YY1=1+2*X1+E1 ; YY2=2+3*X2+E2 ; YY3=3+4*X3+E3$
```

```
? CREATE DUMMY VARIABLES FOR FULL OBSERVABILITY

     CREATE    ; FY1=YY1>0; FY2=YY2>0 $

? CREATE DUMMY VARIABLES FOR CENSORED PROBIT

     CREATE    ; CY1=YY1>0 $
     CREATE    ; IF (FY1=1 & FY2=1) CY2=1; (ELSE) CY2=0$

? CREATE DUMMY VARIABLES FOR POIRIER'S PROBIT

     CREATE    ; Y = FY1*FY2 $

? CREATE Y3

     CREATE    ; Y3 = Y>0*YY3 $
```

As you see from the program, YY1 and YY2 amount to y_{1i}^* and y_{2i}^* in the class; FY1 and FY2 amount to y_{1i} and y_{2i} in full observability model; CY1 and CY2 amount to dummies in censored probit model; and Y = FY1 x FY2. You can ignore the existence of YY3. Using BIV1.DB, You can check how important the quality of given information is. In order to read the data set, use the following commands (BIV_READ.LIM):

```
READ ; NVAR=12 ; NOBS=500
     ; FILE = BIV.DB
     ; NAMES = X1,X2,X3,YY1,YY2,YY3,FY1,FY2,CY1,CY2,Y,Y3 $
```

[EXERCISE]Use BIV.DB. The model is given:

$$y_{1i}^* = {}_{11} + {}_{12}x_{1i} + {}_{1i}$$
$$y_{2i}^* = {}_{21} + {}_{22}x_{2i} + {}_{2i}.$$

Do SURE, full observability, censored, Poirier, and Abowd-Farber MLE. Which method generates the best results in terms of efficiency? (You may fail to get a converged MLE result for Poirier's probit. Read your output carefully. If LIMDEP fails to locate MLE, it will says "abnormal exit from iterations."

[EXERCISE]Use MWPSID82.DB. Consider the joint estimation of labor-force participation and employment decisions. For the labor-force participation, the explanatory variables are:

ONE,ED,MINOR,URB,REGS,AGE,EXP,KIDS5,OFINC,UNEMPR.

For the employment decision, the explanatory variables are the same except that AGE is excluded. Estimate the model by the censored probit. Using your estimates, estimate Pr($y_{1i}^* > 0$ and $y_{2i}^* > 0$) at means of regressors, and compute its standard error.

TOBIT

The model is given: $y_i^* = x_i \ + \ {}_i$; ${}_i$ iid N(0, ^{2}).

Basic Command

(1) $y_i = \max(0,y_i^*)$:

```
TOBIT; LHS = Y ; RHS = ONE,X1,X2,... ; MARGIN ; PAR $
```

(2) $y_i = \min(0,y_i^*)$:

```
TOBIT; LHS = Y ; RHS = ONE,X1,X2,... ; UPPER ; MARGIN ; PAR $
```

(3) $y_i = \min(L_i,y_i^*)$

```
TOBIT; LHS = Y ; RHS = ONE,X1,X2,... ; UPPER ; LIMIT = L ; MARGIN
                  ; PAR $
```

NOTE: For truncation model, replace TOBIT by TRUNC.
NOTE: MARGIN computes $E(y_i \ x_i)/ \ x_i$.
NOTE: The estimate of is stored in S. Can retrieve it using the CALC command.

[EXERCISE]Use MWPSID82.DB. Consider the housework supply of <u>employed</u> married women:

$HWORK_i = X_i \ + \ {}_i$,

where X_i includes ONE, ED, URB, AGE, UNEMPR, OFINC, EXP, OCCW, OCCB, TENURE, WRATE. Estimate this model by Tobit. (Some women reported zero hours on homework.)

SIMPLE SELECTION MODEL

Consider the model:

$y_{1i} = x_{1i}{}_{1} + {}_{1i}$; $y_{2i}^{*} = x_{2i}{}_{2} + {}_{2i}$
$y_{2i} = 1$ if $y_{2i}^{*} > 0$; 0 otherwise.
y_{1i} are observed only if $y_{2i}=1$.

The model can be estimated by following commands:

```
PROBIT; LHS = Y2 ; RHS = list for x2i ; HOLD (IMR = LAM) $
SELECT; LHS = Y1 ; RHS = list for x1i  $
```

This program returns result for the two-stage estimation. B contains ${}_{1}$ and where $= {}_{12}$; VARB contains the corrected covariance matrix. LIMDEP also calculates $= {}_{12}/{}_{1}$ and store it in the name of RHO.

If you want to do MLE, then, use:

```
SELECT; LHS = Y1 ; RHS = list for x1i ; MLE $
```

[EXERCISE]Consider the market wage equation of married women. The model is given:

$LRATE_i = X_{1i}{}_{1} + {}_{1i}$
$EMP_i^{*} = X_{2i}{}_{2} + {}_{2i}$,

where EMP_i^{*} denotes the latent variable which represents the i'th woman's willingness to work. We assume that the woman can get a job immediately if she wants. W_i is observed only if $EMP_i = 1$. Both of X_{1i} and X_{2i} contain ONE, ED, URB, MINOR, AGE, REGS, UNEMPR, LOFINC, KIDS5, EXP. Estimate this model by two-stage and MLE. Can you find some efficiency gains by using MLE?

MOVER/STAYER MODEL

$y_{1i} = x_{1i}\,{}_{1} + {}_{1i},\ {}_{1i}$ N(0, ${}_{1}^{2}$)
$y_{0i} = x_{0i}\,{}_{0} + {}_{0i},\ {}_{0i}$ N(0, ${}_{0}^{2}$)
$I_i^* = z_i\ + u_i.\ u_i$ N(0,1)

${}_{u1} = cov(u, {}_{1});\ {}_{u0} = cov(u, {}_{0})$
${}_{u1} = corr({}_{1},u) = {}_{u1}/{}_{1};\ {}_{u0} = corr({}_{0},u).$

I_i = 1 if I_i^*>0; 0, otherwise.
y_{1i} are observed if I_i = 1.
y_{0i} is observed if I_i = 0.

This model is estimated by following commands:

```
PROBIT     ; LHS = I ; RHS = Z ; HOLD $
SELECT     ; LHS = Y ; RHS = X1 $
MATRIX     ; BETA1 = BSR1 $ ? BSR1 = (b1,s1,rho_{u1})
SELECT     ; LHS = Y ; RHS = X0 ; LIMITS = 1 $
MATRIX     ; BETA0 = BSR0 $ ? BSR0 = (b0,s0,rho_{u2})
SELECT     ; LHS = Y ; RH1 = X1 ; RH2 = X0
           ; MLE ; ALL ; START = BETA1,BETA0
           ; TLF = 0.00001 ; TLB = 0.00001 ; TLG = 0.00001
           ; MAXIT = 1000 $
```

If you run the program, you get following results:

(1) probit .
(2) two-stage ${}_{1},\ {}_{u1}$ (LIMDEP denotes this as coefficient of LAMBDA.)
(3) two-stage ${}_{0},\ {}_{u0}$ (LIMDEP denotes this as coefficient of LAMBDA.)
(4) MLE results: $,\ {}_{1},\ {}_{0},\ {}_{0},\ {}_{u0},\ {}_{1},\ {}_{u1}.$

[EXERCISE]Use MWPSID82.DB. Choose the employed only. You want to estimate union and nonunion wages by LEE's method. In this case, Y is LRATE; all of x_{1i}, x_{2i}, and z_i contain ONE, ED, URB, MINOR, AGE, EXP, REGS, UNEMPR, OCCW, OCCB. Estimate the model by the MOVER/STAYER model. Test H_o: The wage equation is the same for both union and nonunion job. Use a LR test.

MULTIPLE SELECTION MODEL

Consider the double selection model you learned in the class. Assume that y_{1i} and y_{2i} are observed. (Full Observability Model.) The model can be estimated by following commands (MW_SEL1.LIM):

```
PROBIT    ; LHS = Y1 ; RHS = list for X1 $
MATRIX    ; BA = B $
PROBIT    ; LHS = Y2 ; RHS = list for X2 $
MATRIX    ; BB = B $
BIVA  ; LHS = Y1,Y2 ; RH1 = list for X1 ; RH2 = list for X2
          ; START = BA,BB,0 ; HOLD $
SELE  ; LHS = Y3; RHS = list for X3 $
```

Note: You can use full-observability, censored or Poirior probits
Note: I'm not sure this program is correct.

Here is a sample program which computes two-stage estimates using bivariate probit results: (MW_SELE.LIM)

```
NAMELIST; X1=ONE,ED,MINOR,URB,REGS,AGE,EXP,KIDS5,OFINC,UNEMPR
        ; X2=ONE,ED,MINOR,URB,REGS,    EXP,KIDS5,OFINC,UNEMPR $

REJECT  ; LF # 1 $
PROBIT  ; LHS = EMP
        ; RHS = X2 $
MATRIX  ; BP = B   $
SAMPLE  ; ALL $
PROBIT  ; LHS = LF
        ; RHS = X1 $
MATRIX  ; BT  = B  $
BIVARIATE ; LHS = EMP,LF
          ; RH1 = X2; RH2 = X1
          ; START = BP,BT,0 ; TLF = 0.1E-15
          ; SELECTION
          ; PAR
          ; MAXIT = 1000 $

MATRIX  ; BIVCOV = VARB
        ; BIVB   = B
        ; B1 = PART(B,10,19)
        ; B2 = PART(B,1,9)   $
CALC    ; RHOU = RHO         $

? CHOOSE THE GROUP WITH OBSERVED Y3.
? YOU HAVE TO GIVE SELECTION VARIABLE.

REJECT; EMP # 1 $

? GENERATE SELECTIVITY REGRESSORS
? B1,B2 COME FROM BIVARIATE PROBIT.
```

```
? X1 IS FOR FIRST SELECTION RULE, AND X2 FOR SECOND.

CREATE; Z1 = B1'X1
      ; Z2 = B2'X2 $

NAMELIST; Z = Z1,Z2 $

CALC  ; PII = PI $
CREATE; H1 = (Z2-RHOU*Z1)/SQR(1-RHOU^2)
      ; H2 = (Z1-RHOU*Z2)/SQR(1-RHOU^2)
      ; S1 = N01(Z1)*PHI(H1)/BVN(Z,RHOU)
      ; S2 = N01(Z2)*PHI(H2)/BVN(Z,RHOU)
      ; S3 = 1/(2*PII*SQR(1-RHOU^2))
             *EXP( -0.5*(Z1^2-2*RHOU*Z1*Z2+Z2^2)/(1-RHOU^2) )
             /BVN(Z,RHOU) $

? DEFINE REGRESSORS IN Y3

NAMELIST; X = ONE, ED,  MINOR, URB, REGS,
              AGE, EXP, UNEMPR           $

? DEFINE # OF REGRESSORS IN Y3

CALC ; KX = COL(X) $

? DEFINE SELECTIVITY REGRESSORS

NAMELIST; SELX = X, S1, S2 $

? OLS REGRESSION WITH SELECTIVITY REGRESSORS

REGRESS; LHS = LRATE; RHS= SELX; RES = E $

MATRIX; NOLIST
      ; SELB = B $

? CALCULATING CORRECTED VARIANCE OF ERRORS IN EQ. 3
? DEFINE SIG13 AND SIG23. THEY ARE THE LAST TWO
? ENTRIES OF SELB.

CALC  ; C1 = SELB(KX+1)
      ; C2 = SELB(KX+2) $

CREATE; SSSIG33 =  C1*C1*(Z1*S1+S1^2+RHOU*S3)
                 + C2*C2*(Z2*S2+S2^2+RHOU*S3)
                 - 2*C1*C2*(S3-S1*S2)
      ; EE = E^2$
TYPE; CORRECTED VARIANCE OF ERROR IN THIRD EQUATION$

CALC  ; LIST
      ; SIG33  = XBR(SSSIG33) + XBR(EE)$

CREATE; KK =  SIG33-SSSIG33 $
```

```
MATRIX; SELMAT1 = SELX'[KK]SELX $

DELETE; SSSIG33,KK,EE,E$

? GENERATE THE CORRECT COVARIANCE MATRIX OF ERRORS

CALC  ; FO = 1/SQR(1-RHOU^2)$

CREATE; WT1 = C1*(-Z1*S1-S1^2-RHOU*S3) + C2*(S3-S1*S2)
      ; WT2 = C1*(S3-S1*S2) + C2*(-Z2*S2-S2^2-RHOU*S3)
      ; WT3 = C1*(-FO*H2*S3-S1*S3) + C2*(-FO*H1*S3-S2*S3)$

MATRIX; SMAT21   = X1'[WT1]SELX
      ; SMAT22   = X2'[WT2]SELX
      ; SSMAT23  = SELX'[WT3]ONE
      ; SMAT23   = SSMAT23'
      ; SELMAT2  = [SMAT22 / SMAT21 / SMAT23] $

TYPE    ;  CORRECTED  TWO-STAGE  ESTIMATION  RESULT  FOT  DOUBLE
SELECTIVITY$

MATRIX; SELMAT = SELMAT1 + SELMAT2'*BIVCOV*SELMAT2
      ; CORCOV = <SELX'SELX>*SELMAT*<SELX'SELX> $

MATRIX; LIST
      ; STAT(SELB,CORCOV)$

SAMPLE;ALL$
CLOSE
```

If you run this program correctly, you will get:

```
  Variable   Coefficient    Standard  Error   z=b/s.e.  P[  Z   z]
Mean of X

  SELB _ 1   0.49166         0.14487        3.394   0.00069
  SELB _ 2   0.74413E-01     0.69689E-02   10.678   0.00000
     :         :            :        :        :
     :         :            :        :        :
```

[EXERCISE]

1) Explain line by line what the above program does.
2) Run the above program.
3) Estimate the same model using the SELE command. Compare the results with those from 2).
4) Test the hypothesis of no selection by LM statistic.

PANEL DATA MODELS:

Data Configuration

A panel data set is usually arranged in blocks of observations T for a given i. (Your data set may be arranged in blocks of observation N for a given t. In this case, see LIMDEP manual.) Before estimation, you have to give a number for each cross-sectional units and specify a period for each observation in i'th cross-sectional unit. For this, do following:

```
CREATE; IND = TRN(T,0) ; PER = TRN(-T,0) $
```

where T is the time periods in your data set. (T=1,2,3,...) IND is a variable which indicates a specific integer number for a cross-sectional unit, and PER is the variable which indicates periods. LIMDEP's panel data estimator requires a sequential, complete set of indices for the stratification. (Especially, IND is most needed.) Once you do this, you are ready.

Basic Command for Static Model

The basic command for a static model is given:

```
REGRESS    ; LHS = dependent variable
           ; RHS = list of regressors
           ; STR = IND
           ; FIXED ; PANEL ; OUTPUT = 4 $
```

The above command will generate within estimator assuming that the individual effects are fixed effects. **You should not include any time-invariant regressor**. Intercept term is regarded as a time-invariant regressor, and therefore, you should not include ONE either. If you want to estimate the model by GLS, replace ;FIXED by ;RANDOM. When you use random-effects model, you can also estimate the coefficients of time-invariant regressors. Therefore, you may include any regressors you want in the list of regressors. For GLS, **you again should not include ONE.** If you do, LIMDEP will fail to give GLS result. This is because LIMDEP automatically include intercept term for GLS estimation.

If you omit both ;FIXED and ;RANDOM, LIMDEP computes both within and GLS estimators. It also report Hausman test result. This may be a good choice only if your model does not contain any time-invariant regressor other than the overall intercept term. If your model includes time-invariant regressors, the estimation results should be questioned.

The subcommand, ;OUTPUT = 4, is not necessary. I advise you to use it, because it will send a lot of information to the output file.

Outcomes from Estimation of Static Model

For the Case of Fixed-Effects Model

LIMDEP estimates 4 different models:

1) y_{it} = + $_{it}$ (no regressor and no individual effects)
2) y_{it} = $_{i}$ + $_{it}$ (individual effects but no regressors)
3) y_{it} = x_{it} + + $_{it}$ (regressors but no individual effects)
4) y_{it} = x_{it} + $_{i}$ + $_{it}$ (fixed-effects model)

First, LIMDEP will report a table of means for individual observations. (If you have more than 20 individuals, LIMDEP will not show these results on the screen. But those results will be sent to the output file.) The first row of the table is the overall, full sample means, denoted by group '0'.

Second, LIMDEP reports the results for the model 3) above. Output consists of the standard results for OLS regression.

Third, LIMDEP reports results for the model 4), the within estimation results. The estimates of individual effects and standard errors will be listed in your output file.

Fourth, LIMDEP will report some LR and F statistics for the restrictions of:

Model 1) vs. Model 2)
Model 1) vs. Model 3)
Model 1) vs. Model 4)
Model 2) vs. Model 4)
Model 3) vs. Model 4)

The first four statistics are not interesting. The last test statistic is somewhat interesting. If the statistic is small (accepting hypothesis), it implies that all the individual effects are the same. This means that you need only overall intercept term.) If the statistic is large, it means that individual effects differ across i.

Fifth, LIMDEP will store, in the name of SUMSQDEV, SSE (sum of squared residuals) from within estimation. You will need this to make programs for the IV estimation of Hausman-Taylor models. Therefore, <u>DON'T FORGET THIS</u>!!!

For the Case of Random-Effects Model

First, LIMDEP will report estimates of ^{2}(1/T)+ ^{2}, ^{2}, ^{2}, and = corr($_{i}$+ $_{it}$, $_{i}$+ $_{is}$) = ^{2}/(^{2}+ ^{2}).

Second, LIMDEP will report Breusch and Pagan's LM statistic for testing random-effects model against Model 3) above. Large value of the LM statistic argues in favor of random-effects model.

Third, LIMDEP will give you GLS estimates of coefficients of regressors.

Information on PANEL.DB

You can read the data set by the following command: (PAN_READ.LIM)

```
? THIS PROGRAM IS FOR READING PANEL.DB

READ ; NOBS= 2500 ; NVAR=33
     ; FILE=PANEL.DB
     ; FORMAT=(6G13.6)
     ; NAMES=  Y,X1,X2,X3,X4,Z2,Z3,Z4,
          YB,X1B,X2B,X3B,X4B,
          X11AM,X12AM,X13AM,X14AM,X15AM,
          X21AM,X22AM,X23AM,X24AM,X25AM,
          X31AM,X32AM,X33AM,X34AM,X35AM,
          X41AM,X42AM,X43AM,X44AM,X45AM  $
```

Y is the dependent variable; X1,X2,X3, X4 are time-varying regressors; Z2, Z3 and Z4 are time-invariant regressors; X1B, X2B, X3B, X4B and YB are repeated means of corresponding variables (they correspond to $_1$, $_2$, $_3$, $_4$ and in the class); and X11AM is a vector repeated observations of X1 for t =1 (others are similarly defined). There are 500 cross-sectional units with 5 periods. Therefore, there are in total 2500 observations. Before you start estimation, you have to create stratification variables as I explained above.

The following is my program for within, GLS, Hausman test, ALT1, ALT2 and ALT3 (PAN_GLS.LIM):

```
CREATE; IND = TRN(5,0) ; PER = TRN(-5,0) $

? SPECIFIY # OF OBS.
CALC  ; NN = 500 ; TT = 5  $

? SPECIFY # OF TIME-VARYING VARIABLES (KK)
? SPECIFYING # OF TIME-INVARIANT VARIABLES

CALC  ; KK = 4 ; GG = 4 $

? PROGRAMS FOR WITHIN ESTIMATION

CREATE ; WX1 = X1 - X1B
       ; WX2 = X2 - X2B
       ; WX3 = X3 - X3B
       ; WX4 = X4 - X4B
       ; WY  = Y  - YB            $

NAMELIST ; WX = WX1,WX2,WX3,WX4   $
```

```
TYPE     ; WITHIN ESTIMATION RESULTS    $
REGRESS  ; LHS = WY ; RHS = WX          $
CALC     ; LIST ; SIGMAE2 = SUMSQDEV/(NN*(TT-1)-KK)  $

MATRIX  ; WB = B
        ; WCOV = SIGMAE2*XPXI(WX)
        ; STAT(WB,WCOV)                     $

? PROGRAM FOR GLS MODEL

TYPE   ; BETWEEN ESTIMATION RESULTS $
REGRESS; LHS = YB ; RHS = X1B,X2B,X3B,X4B,ONE,Z2,Z3,Z4 $
CALC   ; LIST
       ; SIGMAA2 = SUMSQDEV/(NN-KK-GG)
       ; THET = SQR(SIGMAE2/SIGMAA2)
       ; THE = 1-THET  $
SAMPLE ; ALL $
CREATE ; THEY = Y - THE*YB
       ; THEONE = THET
       ; THEX1 = X1 - THE*X1B
       ; THEX2 = X2 - THE*X2B
       ; THEX3 = X3 - THE*X3B
       ; THEX4 = X4 - THE*X4B
       ; THEZ2 = THET*Z2
       ; THEZ3 = THET*Z3
       ; THEZ4 = THET*Z4   $

NAMELIST;   THEX=THEX1,THEX2,THEX3,THEX4,THEONE,THEZ2,THEZ3,THEZ4
$

REGRESS; LHS = THEY
       ; RHS = THEX
       ; RES = RESY  $
TYPE   ; GLS RESULTS $
MATRIX ; REB = B
       ; RECOV = SIGMAE2*XPXI(THEX)
       ; STAT(REB,RECOV)       $

? HAUSMAN TEST RESULTS
TYPE   ; HAUSMAN TEST RESULTS  $

MATRIX ; REB13 = PART(REB,1,4) ; RECOV13 = PART(RECOV,1,4,1,4) $
MATRIX ; DIFB = WB~REB13 $
MATRIX ; INVCOV = WCOV~RECOV13  $
MATRIX ; DIFV  = SINV(INVCOV)
       ; HTEST = DIFB'|DIFV|DIFB $
CALC   ; LIST
```

```
        ; HT = HTEST
        ; P_VAL = 1 - CHI(HT,KK) $

? ALTERNATIVES OF HAUSMAN TEST

NAMELIST ; TESTX1=WX1,WX2,WX3,WX4,ONE,Z2,Z3,Z4,X1B,X2B,X3B,X4B $
REGRESS  ; LHS = RESY ; RHS = TESTX1
         ; KEEP = FRESY             $

TYPE     ; ALT1 TEST  $
CALC     ; LIST
         ; ALT1 = TT*NN*RSQRD
         ; P_VAL = 1 - CHI(ALT1,KK) $

TYPE     ; ALT2 TEST $
CALC     ; LIST
         ; ALT2 = DOT(FRESY,FRESY)/SIGMAE2
         ; HT = HT
         ; P_VAL = 1 - CHI(ALT2,KK)  $

NAMELIST ; TESTX2 = WX,ONE,Z2,Z3,Z4,
                    X11AM,X12AM,X13AM,X14AM,X15AM,
                    X21AM,X22AM,X23AM,X24AM,X25AM,
                    X31AM,X32AM,X33AM,X34AM,X35AM,
                    X41AM,X42AM,X43AM,X44AM,X45AM $

TYPE     ; ALT3 TEST $
CALC     ; DD = TT*KK ;
REGRESS  ; LHS = RESY ; RHS = TESTX2 $
CALC     ; LIST
         ; ALT3 = TT*NN*RSQRD
         ; P_VAL = 1 - CHI(ALT3,DD) $
```

[EXERCISE] Use PANEL.DB.

1) Run the above program and report the estimation results.
2) Compare the estimation results from 1) and estimation results you can directly obtain using LIMDEP's panel data commands.
3) Assume that Z2, Z3 and Z4 are correlated with the individual effects while other regressors (including ONE) are not. Make a program for Hausman-Taylor estimation method, and report your results.
4) Test the Hausman-Taylor specification using Hausman test method. (Hint: Your result must be favorable for H-T specification.) Use GINV.
5) Assume the same things as 3). Make a program for AM estimation method, and report your results.
6) Test AM specification using Hausman test method.

Probit for Panel Data

The basic command is:

```
PROBIT ; LHS = ... ; RHS = ... ; PDS = T $
```

For probit, the data must be balanced; T, the number of periods must be the same for each i. Data arrangement is identical to the regression model, with multiple records per i. But, here, data must be arranged in N contiguous blocks of exactly T observations.

SOME USEFUL LIMDEP COMMANDS

MATRIX OPERATION

Data Area

LIMDEP store your data in a form similar to Lotus worksheet file. You can read up to 200 variables, and the observations depending on the ram size of your computer.

Namelist

It is used to define data matrices. In a sense, it is akin to creating usual matrices.

```
NAMELIST   ; X = ONE,ED,URB,MINOR $
PROBIT     ; LHS = EMP ; RHS = X $
```

Sample Selection

When you use subsamples,

```
SAMPLE     ; 1-100, 201-300 $
PROBIT     ; LHS = EMP ; RHS = X $
```

If you want to restore all observations,

```
SAMPLE  ; ALL $
```

Reject

You use this command if you want to eliminate some observations.

REJECT ; LF # 1 $ or REJECT ; LF = 0 $

Then, you can eliminate observations for housewives. If you want to restore all the observations,

SAMPLE ; ALL $

Placing matrix results in scalar

1) A IS 1 x 1 MATRIX.

CALC ; AA = A $

Then, A is stored as a scaler in the name of A.

2) A is 4 x 4, and wish to use the (2,1) entry as a scalar

```
MATRIX     ; AA = A(2,1) $
CALC       ; AAA = AA    $
```

IMPORTANT: Do not use the same name in both CALC and MATRIX.

Matrix Algebra

A. Matrix Alebra

1) A|B = AB
2) A~B = A - B
3) A & B = A + B
4) A' = TRANSPOSE OF A
5) SINV(A) = INVERSE OF A, WHERE A IS SYMMETRIC PD.
6) GINV(A) = INVERSE OF A, WHERE A IS ANY SQUARE MATRIX.

IMPORTANT: I strongly recommend you to use SINV for symmetric and pd matrices. During an estimation procedure, you often create symmetric pd matrices, or your matrix outcomes are usually symmetric pd. If SINV does not work, it implies that you made some mistakes in the program. So, you can indirectly check your program using SINV.

7) NVSM(A,B) = INVERSE OF (A + B)
NVSM(A,-B) = INVERSE OF (A - B)

8) INIT(p,q,r) = p x q MATRIX WHOSE ENTRIES ARE ALL r.
9) STAT(B,VARB) = TABLE SHOWING STATISTICAL RESULTS.

B. FOR DATA MATRICES (ASSUME X AND Z ARE DEFINED BY NAMELIST.)

1) XPXI(X) = INVERSE OF X'X = SUM of $x_i'x_i$.
2) XDOT(X,ONE) = SUM OF x_i'
3) XDOT(X,ONE,U) = SUM OF $u_i x_i'$ WHERE U IS A VARIABLE
4) XDOT(X,ONE,Y) = SUM OF x_i y_i WHERE Y IS A VARIABLE
5) XDOT(X) = X'X = sum of x_i x_i
6) XDOT(X,U) = SUM OF $u_i x_i' x_i$.
7) XDOT(X,Z,U) = SUM OF $u_i x_i' z_i$.
8) ASSUME B IS A VECTOR.
   CREATE; XX = DOT(X,B) $

   This command creates a new variable XX.

(EXAMPLE)

```
NAMELIST   ; X = ONE, ED, URB, MINOR $
REGRESS    ; LHS = LRATE ; RHS = X     $
MATRIX     ; OLSB = B                  $
CREATE     ; FLRATE = DOT(X,OLSB)    $
```

Here, FLRATE is the fitted value of LRATE.

Scalar Multiplication

1) 2*IDEN(2) = $2I_2$
2) 1/2*IDEN(2) = $I_2/2$

Using Calculator Commands in Matrix Commands

```
CALC ; PP = 1-2*PHI(1.96) $
MATRIX     ; PP*XDOT(X)
```

Partitioned Matrices

1)~~~~~C~=~[C1,~C2]~~~ ~~~ C~=~[C1`,`C2]

2)~~~~~C~ = ~[C1~/~ C2]~~~ ~~~ C~=~LEFT \[MATRIX { C1 # C2 } RIGHT \]

3)~~~~~C~=~[C11`,~C12~/~C21`,C22]~~~   C~=~LEFT \[MATRIX { C11 & C12 # C21 & C22 } RIGHT \]

Extracting Part of a Matrix

1) A is a vector. You want to create a matrix with 2-10 entries of A. Then, AA = PART(A,2,10).

2) A is a matrix. You want to create a matrix with 2-5 rows and 3-7 columns of A. Then, AA = PART(A,2,5,3,7).

Editing a Matrix

A is a matrix. You want to replace the (2,4)th entry by 3. Then,

```
MATRIX ; A(2,4) = 3 $
```

FOR MORE DETAILS, SEE THE HANDOUT

CREATING VARIABLES

Creating Dummy Variables

(EXAMPLE 1) Assume X,Y,Z are variables:

```
CREATE ; P = Z > 0 ; ZZ = Z = 1 ; A = X > 0 * PHI(Y) $
```

This command creates 3 dummy variables:

1) P = 1 IF Z > 0, AND 0 IF Z <= 0
2) ZZ = 1 IF Z = 1, AND 0, OTHERWISE.
3) A = PHI(Y) IF X > 0, AND 0 IF X <= 0.

(EXAMPLE 2)

```
CREATE    ; IF (AGE >= 56) AGE56 = 1
          ; IF ( 46 <= AGE &  AGE <= 55) AGE46 = 1
          ; IF ( 31 <= AGE &  AGE <= 45) AGE31 = 1
          ; IF ( 23 <= AGE &  AGE <= 30) AGE23 = 1
          ; IF ( AGE <= 22 )             AGE19 = 1  $
```

Creating Variables

(EXAMPLE) Assume C and D are defined as a scaler (by CALC command).

```
CREATE       ; X12 = X1*X2 ; X23 = X2/X3 ; X11 = 1-C*X1
             ; X33 = (C/D)*X1 $
```

Creating Variable Using a Vector

(EXAMPLE 1) Assume that X is defined by namelist.

```
NAMELIST   ; X = ONE,X1,X2,X3 $
REGRESS    ; LHS = Y ; RHS = X $
MATRIX     ; OLSB = B          $
CREATE     ; YFIT = DOT(X,OLSB)      $
```

This program creates $_i = x_i$. Don't use CREATE; YFIT = X*OLSB.

(EXAMPLE 2)

```
NAMELIST; X = ONE,X1,X2,X3 $
PROBIT  ; LHS = Y ; RHS = X $
MATRIX  ; PROBITB = B $
CREATE  ; SELVAR = N01(DOT(X,PROBITB))/PHI(DOT(X,PROBITB)) $
```

This program creates selectivity regressor for the two-stage estimation of Heckman's simple selection model: $_i = (x_i)/ (x_i)$.

Functions for CREATE

See Handout of transformation.

CREATING STANDARD BIVARIATE NORMAL CDF

Suppose you want to create the variable: $z_i = F(x_{1i},x_{2i},)$ where F is cdf for standard bivariate normal. Assume that   is defined by CALC command in the name of RO. (Do not use the name "RHO". This name is reserved for LIMDEP. See the Handout by Dr. Kim, p.1.) Then,

```
NAMELIST; ZZ = X1,X2 $
CREATE  ; Z = BVN(ZZ,RO) $
```

You may replace RO by any number you wish.