Greene

(1)

THEOREM 17.8 Asymptotic Distribution of the Two-Step MLE [Murphy and Topel (1985)]

If the standard regularity conditions are met for both log-likelihood functions, then the second-step maximum likelihood estimator of $\boldsymbol{\theta}_2$ *is consistent and asymptotically normally distributed with asymptotic covariance matrix*

$$\mathbf{V}_2^* = \frac{1}{n}\left[\mathbf{V}_2 + \mathbf{V}_2[\mathbf{CV}_1\mathbf{C}' - \mathbf{RV}_1\mathbf{C}' - \mathbf{CV}_1\mathbf{R}']\mathbf{V}_2\right],$$

where

$$\mathbf{V}_1 = \text{Asy.Var}[\sqrt{n}(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1)] \text{ based on } \ln L_1,$$

$$\mathbf{V}_2 = \text{Asy.Var}[\sqrt{n}(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2)] \text{ based on } \ln L_2 \mid \boldsymbol{\theta}_1,$$

$$\mathbf{C} = E\left[\frac{1}{n}\left(\frac{\partial \ln L_2}{\partial \boldsymbol{\theta}_2}\right)\left(\frac{\partial \ln L_2}{\partial \boldsymbol{\theta}_1'}\right)\right], \quad \mathbf{R} = E\left[\frac{1}{n}\left(\frac{\partial \ln L_2}{\partial \boldsymbol{\theta}_2}\right)\left(\frac{\partial \ln L_1}{\partial \boldsymbol{\theta}_1'}\right)\right].$$

The correction of the asymptotic covariance matrix at the second step requires some additional computation. Matrices $\mathbf{V}_1$ *and* $\mathbf{V}_2$ *are estimated by the respective uncorrected covariance matrices. Typically, the BHHH estimators,*

$$\hat{\mathbf{V}}_1 = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \ln f_{i1}}{\partial \hat{\boldsymbol{\theta}}_1}\right)\left(\frac{\partial \ln f_{i1}}{\partial \hat{\boldsymbol{\theta}}_1'}\right)\right]^{-1}$$

and

$$\hat{\mathbf{V}}_2 = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \ln f_{i2}}{\partial \hat{\boldsymbol{\theta}}_2}\right)\left(\frac{\partial \ln f_{i2}}{\partial \hat{\boldsymbol{\theta}}_2'}\right)\right]^{-1}$$

are used. The matrices $\mathbf{R}$ *and* $\mathbf{C}$ *are obtained by summing the individual observations on the cross products of the derivatives. These are estimated with*

$$\hat{\mathbf{C}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \ln f_{i2}}{\partial \hat{\boldsymbol{\theta}}_2}\right)\left(\frac{\partial \ln f_{i2}}{\partial \hat{\boldsymbol{\theta}}_1'}\right)$$

and

$$\hat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \ln f_{i2}}{\partial \hat{\boldsymbol{\theta}}_2}\right)\left(\frac{\partial \ln f_{i1}}{\partial \hat{\boldsymbol{\theta}}_1'}\right)$$

Example 17.9 Two-Step ML Estimation

Continuing the example discussed at the beginning of this section, we suppose that y_{i2} is a binary indicator of the choice whether to enroll in the program ($y_{i2} = 1$) or not ($y_{i2} = 0$) and that the probabilities of the two outcomes are

$$\text{Prob}[y_{i2} = 1 \mid \mathbf{x}_{i1}, \mathbf{x}_{i2}] = \frac{e^{\mathbf{x}_{i2}'\beta + \gamma E[y_{i1} \mid \mathbf{x}_{i1}']}}{1 + e^{\mathbf{x}_{i2}'\beta + \gamma E[y_{i1} \mid \mathbf{x}_{i1}']}}$$

1. STATA (Jay's note on Stata Codes)

This file is intended to help understanding the Murphy correction. I only go through the technical details which I suppose may be hard to be understood.

Model: We want to estimate a probit model in which an explanatory variable is the fitted value from the previous linear regression.

Objective: We need to correct the covariance matrix to take into account the fact that one independent variable is generated variable.

Stage one. Run a linear regression and keep the fitted value. The regression is given by:

```
regress hours nwifeinc educ exper expersq age
```

We keep the fitted value as "what" and the residuals as "epi"

```
predict double what
predict double epi, res
```

Since Stata makes some degree-of-freedom adjustment, we need to undo this adjustment to obtain the covariance matrix yielded by MLE.
The corrected covariance is given by v1:

```
matrix v1=(e(df_r)/e(N))*e(V)
```

In addition, we need the MLE estimator for the sigma square (σ^2):

```
scalar sig = e(rss)/e(N)
```

Stage two. Now we run the probit regression by including the fitted value "what" as one of the RHS variables.

```
probit inlf what nwifeinc educ exper expersq age hushrs husage, score(s)
```

Note that we need to preserve the score vector (whose definition is different from that given in textbook) as variable "s".

"s" is an $N \times 1$ vector whose i-th element is defined as:

$s_i = \dfrac{\partial \log f_i}{\partial (x_i'\theta)}$, where f_i is the density function for observation i.

That is, stata treats $x_i'\theta$ as a whole, and then take derivative with respect to this cross product. The link between this "stata-score" and usual "textbook-score" can be established as:

$S^{textbook} = \sum_{i=1}^{N} s_i^{stata} x_i$, where x_i is a $p \times 1$ vector for the observation i.

The advantage of defining this "stata-score" is huge. We can express any "textbook - score vector" as:

$\left(\dfrac{\partial LogL}{\partial \theta_k}\right) = (\sum_{i=1}^{N} s_i x_{ik}), where$ k=1,2,....p, and i=1,2,...N.

Stage three. Now we are ready to do the hard job. Please read the Greene's textbook, page 508-512 for a concise derivation. If you want to see how this complicated problem can be made more complicated, read Hardin's paper.

The key components we need to construct are two matrix of the form of outer products.

The first matrix is estimated as:

$$\hat{C} = N^{-1}\sum[(\partial \ln f_{2i}/\partial\theta_2)(\partial \ln f_{2i}/\partial\theta_1')|_{\theta_2=\hat{\theta}_2,\theta_1=\hat{\theta}_1}]$$

and the second one is given as:

$$\hat{R} = N^{-1}\sum[(\partial \ln f_{2i}/\partial\theta_2)(\partial \ln f_{1i}/\partial\theta_1')|_{\theta_2=\hat{\theta}_2,\theta_1=\hat{\theta}_1}]$$

Now I will make explicit the expression involving partial derivatives.
To reproduce the log-likelihood for the first-stage linear regression for observation i:

$\ln f_{1i} = cons - (1/2)\ln(\sigma^2) - (y_{1i} - x_{1i}'\theta_1)^2/(2\sigma^2)$ (Using normal density)

This implies that

$\{\partial \ln f_{1i}/\partial\theta_1'|_{\theta_1=\hat{\theta}_1}\} = (y_{1i} - x_{1i}'\hat{\theta}_1)x_{1i}'/\sigma^2 = e_{1i}x_{1i}'/\sigma^2$; where e_{1i} is the OLS residual for observation i.

In the following derivation, we need to utilize the definition of "stata-score". First note that $\{\partial \ln f_{2i}/\partial\theta_2|_{\theta_2=\hat{\theta}_2}\} = [\partial \ln f_{2i}/\partial(x_{2i}'\hat{\theta}_2)]*[\partial(x_{2i}'\hat{\theta}_2)/\partial\hat{\theta}_2] = s_i x_{2i}$. Don't bother to derive the explicit form for the derivative. Stata can handle it numerically. All the information we need is just the "stata-score".

In a similar fashion we can show that:

$\{\partial \ln f_{2i}/\partial\theta_1'|_{\theta_1=\hat{\theta}_1}\} = [\partial \ln f_{2i}/\partial(x_{2i}'\hat{\theta}_2)]*[b\_what * x_{1i}'] = b\_what * s_i x_{1i}'$;

where b_what is the coefficient of "what" in the probit estimation.

Regarding the vector calculus, when you take derivative with a column (row) vector, the result should be a column (row) vector. Here $\hat{\theta}$ is a $p\times 1$ column vector and $\hat{\theta}'$ is a $1\times p$ row vector.

Short question: can we define "stata-score" in the first stage linear regression even if it's estimated by OLS? (answer: yes. The "stata-score" is just the normalized residual vector.)

Stage four: Now two thirds of job has been done. We observe that:

$$\sum[(\partial \ln f_{2i}/\partial\theta_2)(\partial \ln f_{2i}/\partial\theta_1')|_{\theta_2=\hat{\theta}_2,\theta_1=\hat{\theta}_1}] = \sum b\_what * s_i^2 x_{2i}x_{1i}'$$

$$= b\_what * (x_{21}, x_{22}, \ldots, x_{2N})diag(s_1^2, \ldots s_N^2)(x_{11}', x_{12}', \ldots, x_{1N}')'$$

$$= X_2'WX_1$$

Where the weighting matrix is define as $W = b\_what * diag(s_1^2, \ldots s_N^2)$, and $diag(\ldots)$ denotes a diagonal matrix.

If you can understand so far, you will have little trouble to derive the R matrix. Try by yourself. In stata, the C and R matrix can be constructed using the command "`matrix accum`" along with the specification of the weighting matrix. The remaining job is just following the instruction of Greene.

4

```
set more off
log using Murphy_topel_log.smcl, replace

* you'd better read the Greene's book, page 510.

*******************************
* Murphy's variance estimator *
*******************************

*insheet using mroz.txt, clear
use mroz.dta

describe

* first stage__linear regression model

regress hours nwifeinc educ exper expersq age

* to undo the correction of degree of freedom, in order to obtain the MLE
variance matrix
matrix v1=(e(df_r)/e(N))*e(V)
predict double what
predict double epi, res
scalar sig = e(rss)/e(N)

* second stage_probit regression without correction to the variance matrix.
* keep the score as "s". This score is different from its normal meaning.

probit inlf what nwifeinc educ exper expersq age hushrs husage, score(s)
* the variable exper is dropped due to collinearity
matrix v2= e(V)
scalar zz = _b[what]
gen ss = s^2

*********************************
* correction of variance matrix *
*********************************
gen byte cons = 1
matrix accum C = what nwifeinc educ expersq age hushrs husage cons nwifeinc
educ exper expersq age cons /*
                  */ [iw=ss*zz], nocons

matrix accum R = what nwifeinc educ expersq age hushrs husage cons nwifeinc
educ exper expersq age cons /*
                  */ [iw=s*epi/sig], nocons

matrix c = C[1..8,9..14]

matrix r = R[1..8,9..14]

matrix m = (v2 + (v2*(c*v1*c'-r*v1*c'-c*v1*r')*v2))

dis "following is the incorrect covariance matrix of theta2"
matrix list v2

dis "following is the correct covariance matrix of theta2"
matrix list m

log close
```

```
  ___  ____  ____  ____  ____tm
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/
  Statistics/Data Analysis
```

```
      log:  C:\Documents and Settings\jlee\My Documents\Teaching\EC671\Murphy_topel_log.smcl
 log type:  smcl
opened on:  21 Sep 2007, 11:05:39

 1 .
 2 . * you'd better read the Greene's book, page 510.
 3 .
 4 . *******************************
 5 . * Murphy's variance estimator *
 6 . *******************************
 7 .
 8 . *insheet using mroz.txt, clear
 9 . use mroz.dta

10 .
11 . describe

Contains data from mroz.dta
  obs:           753
 vars:            22                          2 Sep 1996 16:04
 size:        39,909 (96.2% of memory free)
```

variable name	storage type	display format	value label	variable label
inlf	byte	%9.0g		=1 if in lab frce, 1975
hours	int	%9.0g		hours worked, 1975
kidslt6	byte	%9.0g		# kids < 6 years
kidsge6	byte	%9.0g		# kids 6-18
age	byte	%9.0g		woman's age in yrs
educ	byte	%9.0g		years of schooling
wage	float	%9.0g		est. wage from earn, hrs
repwage	float	%9.0g		rep. wage at interview in 1976
hushrs	int	%9.0g		hours worked by husband, 1975
husage	byte	%9.0g		husband's age
huseduc	byte	%9.0g		husband's years of schooling
huswage	float	%9.0g		husband's hourly wage, 1975
faminc	float	%9.0g		family income, 1975
mtr	float	%9.0g		fed. marg. tax rte facing woman
motheduc	byte	%9.0g		mother's years of schooling
fatheduc	byte	%9.0g		father's years of schooling
unem	float	%9.0g		unem. rate in county of resid.
city	byte	%9.0g		=1 if live in SMSA
exper	byte	%9.0g		actual labor mkt exper
nwifeinc	float	%9.0g		(faminc - wage*hours)/1000
lwage	float	%9.0g		log(wage)
expersq	int	%9.0g		exper^2

```
Sorted by:

12 .
13 . * first stage__linear regression model
14 .
15 . regress hours nwifeinc educ exper expersq age
```

Source	SS	df	MS
Model	119690318	5	23938063.6
Residual	451219406	747	604042.043
Total	570909724	752	759188.463

Number of obs =	753
F(5, 747) =	39.63
Prob > F =	0.0000
R-squared =	0.2096
Adj R-squared =	0.2044
Root MSE =	777.2

hours	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
nwifeinc	-4.40392	2.630143	-1.67	0.094	-9.567272	.7594327
educ	24.75008	13.33234	1.86	0.064	-1.423234	50.9234
exper	74.51011	10.23346	7.28	0.000	54.42034	94.59989
expersq	-.9266305	.3347942	-2.77	0.006	-1.58388	-.2693811
age	-16.34061	3.902102	-4.19	0.000	-24.001	-8.680215
_cons	593.0894	240.2188	2.47	0.014	121.5051	1064.674

```
16 .
17 . * to undo the correction of degree of freedom, in order to obtain the MLE variance matrix
18 . matrix v1=(e(df_r)/e(N))*e(V)

19 . predict double what
   (option xb assumed; fitted values)

20 . predict double epi, res

21 . scalar sig = e(rss)/e(N)

22 .
23 . * second stage_probit regression without correction to the variance matrix.
24 . * keep the score as "s". This score is different from its normal meaning.
25 .
26 . probit inlf what nwifeinc educ exper expersq age hushrs husage, score(s)

   note: exper dropped due to collinearity
   Iteration 0:   log likelihood =  -514.8732
   Iteration 1:   log likelihood = -433.56257
   Iteration 2:   log likelihood = -431.93358
   Iteration 3:   log likelihood = -431.93128

   Probit estimates                                  Number of obs   =        753
                                                     LR chi2(7)      =     165.88
                                                     Prob > chi2     =     0.0000
   Log likelihood = -431.93128                       Pseudo R2       =     0.1611
```

inlf	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
what	.00166	.0002449	6.78	0.000	.00118	.0021401
nwifeinc	-.0032206	.0048412	-0.67	0.506	-.0127093	.0062681
educ	.0663945	.0252791	2.63	0.009	.0168485	.1159406
expersq	-.0004359	.0003953	-1.10	0.270	-.0012106	.0003388
age	-.010028	.0142904	-0.70	0.483	-.0380367	.0179807
hushrs	-.0000949	.0000845	-1.12	0.262	-.0002606	.0000708
husage	.0086198	.0133027	0.65	0.517	-.0174531	.0346926
_cons	-1.454477	.5198501	-2.80	0.005	-2.473364	-.4355891

```
27 . * the variable exper is dropped due to collinearity
28 . matrix v2= e(V)

29 . scalar zz = _b[what]
```

```
30 . gen ss = s^2

31 .
32 . *********************************
33 . * correction of variance matrix *
34 . *********************************
35 . gen byte cons = 1

36 . matrix accum C = what nwifeinc educ expersq age hushrs husage cons nwifeinc educ exper expersq age cc
   >  /*
   >                  */ [iw=ss*zz], nocons
   (obs=.6802719787)

37 .
38 . matrix accum R = what nwifeinc educ expersq age hushrs husage cons nwifeinc educ exper expersq age cc
   >  /*
   >                  */ [iw=s*epi/sig], nocons
   (obs=.4896954242)

39 .
40 . matrix c = C[1..8,9..14]

41 .
42 . matrix r = R[1..8,9..14]

43 .
44 . matrix m = (v2 + (v2*(c*v1*c'-r*v1*c'-c*v1*r')*v2))

45 .
46 . dis "following is the incorrect covariance matrix of theta2"
   following is the incorrect covariance matrix of theta2

47 . matrix list v2

   symmetric v2[8,8]
                 what    nwifeinc        educ     expersq         age      hushrs      husage        _coı
       what  5.998e-08
   nwifeinc  3.251e-07   .00002344
       educ -1.924e-06  -.00004644   .00063903
    expersq -7.804e-08  -2.028e-07   2.143e-06   1.562e-07
        age  9.333e-07  -9.565e-07  -.00002314  -2.127e-06   .00020422
     hushrs  1.374e-09  -4.808e-08  -1.475e-07  -2.162e-09   3.201e-08   7.148e-09
     husage -3.649e-08  -8.081e-08   .00001844   3.077e-07  -.00015929   5.295e-08   .00017696
      _cons -.00005475   .00004215  -.00534087   .00009009  -.00158955  -.00001783  -.00157243   .2702441

48 .
49 . dis "following is the correct covariance matrix of theta2"
   following is the correct covariance matrix of theta2

50 . matrix list m

   symmetric m[8,8]
                 what    nwifeinc        educ     expersq         age      hushrs      husage        _coı
       what  2.763e-08
   nwifeinc  1.076e-07   .00001838
       educ -3.251e-07    -.000024     .000325
    expersq -2.342e-08   7.107e-08   2.562e-08   4.022e-08
        age  3.655e-07  -3.777e-06  -9.473e-06  -1.053e-06   .00018122
     hushrs  7.435e-10  -6.560e-08  -1.639e-07  -1.580e-09   4.560e-09   7.157e-09
     husage -4.842e-08   3.640e-07   .00001073   3.441e-07  -.00016057   4.904e-08   .00017699
      _cons -.00003146   .00010834   -.0027951   .00004261  -.00039394  -.00001559  -.00140991   .164605

51 .
52 . log close
         log:  C:\Documents and Settings\jlee\My Documents\Teaching\EC671\Murphy_topel_log.smcl
    log type:  smcl
   closed on:  21 Sep 2007, 11:05:39
```

2_LIMDEP


```
? Murphy & Topell procedure with an example

? MROZ.dat, Wooldridge, p. 544


Read  ; Nobs = 753
        ; Nvar = 22
        ; file = mroz.dat
      ; names = 3 $

Namelist ; X1 = nwifeinc, educ, exper, expersq, age$
namelist ; xx1= nwifeinc, educ, expersq, age $

? TOBIT WITH SAMPLE SELECTION
 Probit ; lhs = inlf
          ; rhs = one, x1, hushrs, husage
          ; hold $
 Select ; Tobit ; MLE
          ; lhs = hours
          ; rhs = one, x1 $

?=====================================================================
? Two step estimation of a probit model when one of the RHS variables
? is the prediction from a first step linear regression model.
? Computations are based on Theorem 4.19 in Greene (1997, p. 142) and
? Murphy and Topel (1985).
?
? User must define the following:
? NAMELIST ; X = RHS in the probit model, not including fitted w
?          ; Z = RHS in the regression $
? CREATE   ; y = LHS in the probit model

?          ; w = LHS in the regression $
?=====================================================================
?
? Fit the regression and keep results.  Use ML rather than degrees of
? freedom corrected least squares results.  Need residuals, preciction,
? asymptotic covariance matrix and disturbance variance estimator.
? V1 is (e'e/N) * <X'X>

NAMELIST ; X = one, x1, hushrs, husage
         ; Z = one, x1 $
CREATE   ; y = inlf
         ; w = hours $

  REGRESS  ; Lhs = w ; Rhs = Z ; Keep = wfit ; Res = ew $
  MATRIX   ; V1  = {(N-Kreg)/N} * Varb $ (Use ML)
  CALC     ; sigmasq = sumsqdev/N $

?
? Now, fit probit model and correct asymptotic covariance matrix.
? Lambda is used to compute derivatives of log-likelihood.  Note
? dLog-L/dbeta = X'Lambda.
?
  NAMELIST ; XW = wfit,xx1,hushrs,husage,one $
  PROBIT   ; Lhs = y ; Rhs = XW ; Hold(IMR = lambda) $
?
? Two matrices C and R are of the form XW'[variable]Z where variable
? is computed from derivatives of log-likelihoods.  The V2 matrix in
? the computation is the asymptotic covariance matrix from the probit
? model, before the correction.

?
  CREATE   ; vc = B(1) * lambda^2
           ; vr = lambda * ew/sigmasq $
  MATRIX   ; C = XW'[vc]Z
           ; R = XW'[vr]Z
           ; V = C * V1 * C' - R * V1 * C' - C * V1 * R'
```

(9)

```
          ; V = Varb + Varb * V * Varb
          ; stat(B,V) $

  matrix; list; v $

? We get the same results given by Stata.
```

(10)

```
--> RESET
--> Read  ; Nobs = 753
        ; Nvar = 22
        ; file = mroz.dat
    ; names = 3 $
--> Namelist ; X1 = nwifeinc, educ, exper, expersq, age$
--> namelist ; xx1= nwifeinc, educ, expersq, age $
--> Probit ; lhs = inlf
           ; rhs = one, x1, hushrs, husage
           ; hold $
Normal exit from iterations. Exit status=0.

+---------------------------------------------+
| Binomial Probit Model                       |
| Maximum Likelihood Estimates                |
| Model estimated: Sep 21, 2007 at 11:23:16AM.|
| Dependent variable                INLF      |
| Weighting variable                None      |
| Number of observations             753      |
| Iterations completed                 5      |
| Log likelihood function      -431.9313      |
| Restricted log likelihood    -514.8732      |
| Chi squared                   165.8838      |
| Degrees of freedom                   7      |
| Prob[ChiSqd > value] =        .0000000      |
| Results retained for SELECTION model.       |
| Hosmer-Lemeshow chi-squared =  10.43597     |
| P-value=  .23575 with deg.fr. =       8     |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Index function for probability
 Constant    -.4699192654      .47580973     -.988   .3233
 NWIFEINC -.1053131393E-01  .46623263E-02   -2.259   .0239     20.128964
 EDUC        .1074808684    .24094759E-01    4.461   .0000     12.286853
 EXPER       .1236904330    .18248173E-01    6.778   .0000     10.630810
 EXPERSQ  -.1974154961E-02  .59360229E-03   -3.326   .0009     178.03851
 AGE      -.3715424633E-01  .13774305E-01   -2.697   .0070     42.537849
 HUSHRS   -.9488560213E-04  .84547111E-04   -1.122   .2617     2267.2709
 HUSAGE    .8619754463E-02  .13302745E-01     .648   .5170     45.120850
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)


+----------------------------------------+
| Fit Measures for Binomial Choice Model |
| Probit   model for variable INLF       |
+----------------------------------------+
| Proportions P0= .431607   P1= .568393  |
| N =      753 N0=     325   N1=     428 |
| LogL =  -431.93128 LogL0 =  -514.8732  |
| Estrella = 1-(L/L0)^(-2L0/n) = .21354  |
+----------------------------------------+
|     Efron |  McFadden  |  Ben./Lerman  |
|    .20348 |    .16109  |       .60943  |
|    Cramer | Veall/Zim. |     Rsqrd_ML  |
|    .20392 |    .31254  |       .19772  |
+----------------------------------------+
| Information  Akaike I.C. Schwarz I.C.  |
| Criteria        1.16848    916.85509   |
+----------------------------------------+
Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.
Threshold value for predicting Y=1 = .5000
            Predicted
------  ----------  +  -----
Actual      0    1  |  Total
------  ----------  +  -----
```

(11)

```
  0        189  136  |    325
  1         96  332  |    428
------  ----------  +  -----
Total      285  468  |    753
--> Select ; Tobit ; MLE
          ; lhs = hours
          ; rhs = one, x1 $
          +------------------------------------------------------------+
          | Sample Selection Model                                     |
          | Probit selection equation based on INLF                    |
          | Selection rule is: Observations with INLF     =  1         |
          | Results of selection:                                      |
          |                       Data points     Sum of weights       |
          | Data set                  753              753.0           |
          | Selected sample           428              428.0           |
          +------------------------------------------------------------+

+-----------------------------------------------------------------------+
| Sample Selection Model                                                |
| Two stage   least squares regression    Weighting variable = none     |
| Dep. var. = HOURS    Mean=   1302.929907    , S.D.=   776.2743846     |
| Model size: Observations =      428, Parameters =   7, Deg.Fr.=   421 |
| Residuals:  Sum of squares= 225378864.8     , Std.Dev.=     731.67050 |
| Fit:        R-squared=  .109536, Adjusted R-squared =          .09685 |
|             (Note:  Not using OLS.  R-squared is not bounded in [0,1] |
| Model test: F[  6,     421] =    8.63,     Prob value =        .00000 |
| Diagnostic: Log-L =  -3426.5781, Restricted(b=0) Log-L =   -3454.9337 |
|             LogAmemiyaPrCrt.=   13.207, Akaike Info. Crt.=     16.045 |
| Standard error corrected for selection.....    733.90                 |
| Correlation of disturbance in regression                              |
| and Selection Criterion (Rho).............     .10865                 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant    1401.134906       874.99266    1.601   .1093
 NWIFEINC   -.2320387237       5.7881986    -.040   .9680     18.937483
 EDUC       -19.77405191       42.190155    -.469   .6393     12.658879
 EXPER       61.49385752       53.176344    1.156   .2475     13.037383
 EXPERSQ    -.7848854757       .98543429    -.796   .4257     234.71963
 AGE        -12.07120966       12.036581   -1.003   .3159     41.971963
 LAMBDA      79.73960772       701.90590     .114   .9096     .57271354

Normal exit from iterations. Exit status=0.

+---------------------------------------------+
| ML Estimates of Selection Model             |
| Maximum Likelihood Estimates                |
| Model estimated: Sep 21, 2007 at 11:23:17AM.|
| Dependent variable                 HOURS    |
| Weighting variable                  None    |
| Number of observations               753    |
| Iterations completed                  44    |
| Log likelihood function        -3859.237    |
| LHS is CENSORED. Tobit Model fit by MLE.    |
| FIRST  8 estimates are probit equation.     |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
          Selection (probit) equation for INLF
 Constant    -.5330649801      .47360553   -1.126   .2604
 NWIFEINC -.1091377384E-01  .42958412E-02  -2.541   .0111
 EDUC        .1049685724    .25173470E-01   4.170   .0000
 EXPER       .1204931180    .19013083E-01   6.337   .0000
 EXPERSQ  -.1915935570E-02  .66713672E-03  -2.872   .0041
 AGE      -.3243802112E-01  .12839726E-01  -2.526   .0115
```

(2)

```
 HUSHRS   -.5089743216E-04  .69736891E-04    -.730   .4655
 HUSAGE    .4746002466E-02  .12068069E-01     .393   .6941
          Corrected regression, Regime 1
 Constant      .7442859083      273.73143     .003   .9978
 NWIFEINC -.4702368524E-02  .21194178E-01    -.222   .8244
 EDUC      .1760113906E-01      .35242274     .050   .9602
 EXPER         .1228609660      2.0309019     .060   .9518
 EXPERSQ  -.1820095748E-02  .96702804E-03   -1.882   .0598
 AGE      -.2501085894E-01      .15774501    -.159   .8740
 SIGMA(1)  .1121378468E-02  .52855157E-04   21.216   .0000
 RHO(1,2)      .6231884769  .45967633E-01   13.557   .0000
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)


--> NAMELIST ; X = one, x1, hushrs, husage
    ; Z = one, x1 $
--> CREATE   ; y = inlf
    ; w = hours $
--> REGRESS  ; Lhs = w ; Rhs = Z ; Keep = wfit ; Res = ew $

+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = W         Mean=   740.5763612    , S.D.=    871.3142158   |
| Model size: Observations =      753, Parameters =    6, Deg.Fr.=   747 |
| Residuals:  Sum of squares= 451219406.2     , Std.Dev.=     777.20142 |
| Fit:        R-squared=  .209648, Adjusted R-squared =          .20436 |
| Model test: F[  5,     747] =    39.63,    Prob value =         .00000 |
| Diagnostic: Log-L =   -6077.1905, Restricted(b=0) Log-L =  -6165.7724 |
|             LogAmemiyaPrCrt.=   13.319, Akaike Info. Crt.=     16.157 |
| Autocorrel: Durbin-Watson Statistic =    1.26131,    Rho =     .36934 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant     593.0894036       240.21881    2.469   .0136
 NWIFEINC    -4.403919391       2.6301435   -1.674   .0941    20.128964
 EDUC         24.75008246       13.332342    1.856   .0634    12.286853
 EXPER        74.51011363       10.233464    7.281   .0000    10.630810
 EXPERSQ     -.9266304862       .33479416   -2.768   .0056    178.03851
 AGE         -16.34060727       3.9021023   -4.188   .0000    42.537849


--> MATRIX   ; V1  = {(N-Kreg)/N} * Varb $ (Use ML)
--> CALC     ; sigmasq = sumsqdev/N $
--> NAMELIST ; XW = wfit,xx1,hushrs,husage,one $
--> PROBIT   ; Lhs = y ; Rhs = XW ; Hold(IMR = lambda) $
Normal exit from iterations. Exit status=0.

+---------------------------------------------+
| Binomial Probit Model                       |
| Maximum Likelihood Estimates                |
| Model estimated: Sep 21, 2007 at 11:23:18AM.|
| Dependent variable                  Y       |
| Weighting variable               None       |
| Number of observations            753       |
| Iterations completed                5       |
| Log likelihood function     -431.9313       |
| Restricted log likelihood   -514.8732       |
| Chi squared                  165.8838       |
| Degrees of freedom                  7       |
| Prob[ChiSqd > value] =       .0000000       |
| Results retained for SELECTION model.       |
| Hosmer-Lemeshow chi-squared =  10.43597     |
| P-value=  .23575 with deg.fr. =       8     |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
```

(13)

```
          Index function for probability
 WFIT       .1660048911E-02  .24490868E-03    6.778   .0000      740.57636
 NWIFEINC -.3220592342E-02  .48412460E-02    -.665   .5059      20.128964
 EDUC       .6639452094E-01  .25279084E-01    2.626   .0086      12.286853
 EXPERSQ  -.4359030317E-03  .39525359E-03   -1.103   .2701      178.03851
 AGE      -.1002803903E-01  .14290425E-01    -.702   .4828      42.537849
 HUSHRS   -.9488560213E-04  .84547111E-04   -1.122   .2617      2267.2709
 HUSAGE     .8619754463E-02  .13302745E-01     .648   .5170      45.120850
 Constant    -1.454476684     .51985051    -2.798   .0051
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)


+----------------------------------------+
| Fit Measures for Binomial Choice Model |
| Probit   model for variable Y          |
+----------------------------------------+
| Proportions P0= .431607    P1= .568393 |
| N =      753 N0=      325   N1=     428 |
| LogL =   -431.93128 LogL0 =   -514.8732 |
| Estrella = 1-(L/L0)^(-2L0/n) = .21354  |
+----------------------------------------+
|     Efron |  McFadden  |  Ben./Lerman |
|    .20348 |    .16109  |       .60943 |
|    Cramer | Veall/Zim. |      Rsqrd_ML |
|    .20392 |    .31254  |       .19772 |
+----------------------------------------+
| Information  Akaike I.C. Schwarz I.C.  |
| Criteria        1.16848     916.85509  |
+----------------------------------------+
Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.
Threshold value for predicting Y=1 = .5000
              Predicted
------  ----------  +  -----
Actual      0    1  |  Total
------  ----------  +  -----
  0       189  136  |    325
  1        96  332  |    428
------  ----------  +  -----
Total     285  468  |    753
--> CREATE   ; vc = B(1) * lambda^2
    ; vr = lambda * ew/sigmasq $
--> MATRIX   ; C = XW'[vc]Z
    ; R = XW'[vr]Z
    ; V = C * V1 * C' - R * V1 * C' - C * V1 * R'
    ; V = Varb + Varb * V * Varb
    ; stat(B,V) $

+-----------------------------------------------------+
|Number of observations in current sample =      753 |
|Number of parameters computed here       =        8 |
|Number of degrees of freedom             =      745 |
+-----------------------------------------------------+
+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
 B_1       .1660048911E-02  .16620942E-03    9.988   .0000
 B_2     -.3220592342E-02  .42872493E-02    -.751   .4525
 B_3       .6639452094E-01  .18027808E-01    3.683   .0002
 B_4     -.4359030317E-03  .20055144E-03   -2.174   .0297
 B_5     -.1002803903E-01  .13461932E-01    -.745   .4563
 B_6     -.9488560213E-04  .84602002E-04   -1.122   .2621
 B_7       .8619754463E-02  .13303752E-01     .648   .5170
 B_8        -1.454476684     .40571665    -3.585   .0003
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)

--> matrix; list; v $
```

14

2.76256e-008	1.07618e-007	-3.25069e-007	-2.342e-008	3.655e-007
7.43486e-010	-4.84209e-008	-3.14607e-005		
1.07618e-007	1.83805e-005	-2.40047e-005	7.10666e-008	-3.77655e-006
-6.56029e-008	3.64039e-007	0.000108341		
-3.25069e-007	-2.40047e-005	0.000325002	2.5618e-008	-9.47265e-006
-1.63853e-007	1.07256e-005	-0.0027951		
-2.342e-008	7.10666e-008	2.5618e-008	4.02209e-008	-1.0529e-006
-1.5802e-009	3.44078e-007	4.26111e-005		
3.655e-007	-3.77655e-006	-9.47265e-006	-1.0529e-006	0.000181224
4.56032e-009	-0.000160566	-0.00039394		
7.43486e-010	-6.56029e-008	-1.63853e-007	-1.5802e-009	4.56032e-009
7.1575e-009	4.90428e-008	-1.55901e-005		
-4.84209e-008	3.64039e-007	1.07256e-005	3.44078e-007	-0.000160566
4.90428e-008	0.00017699	-0.00140992		
-3.14607e-005	0.000108341	-0.0027951	4.26111e-005	-0.00039394
-1.55901e-005	-0.00140992	0.164606		