Labor market participation: Binary dependent variable

Source: MROZ.RAW in Wooldridge. T.A. Mroz (1987), "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica* 55, 765-799.

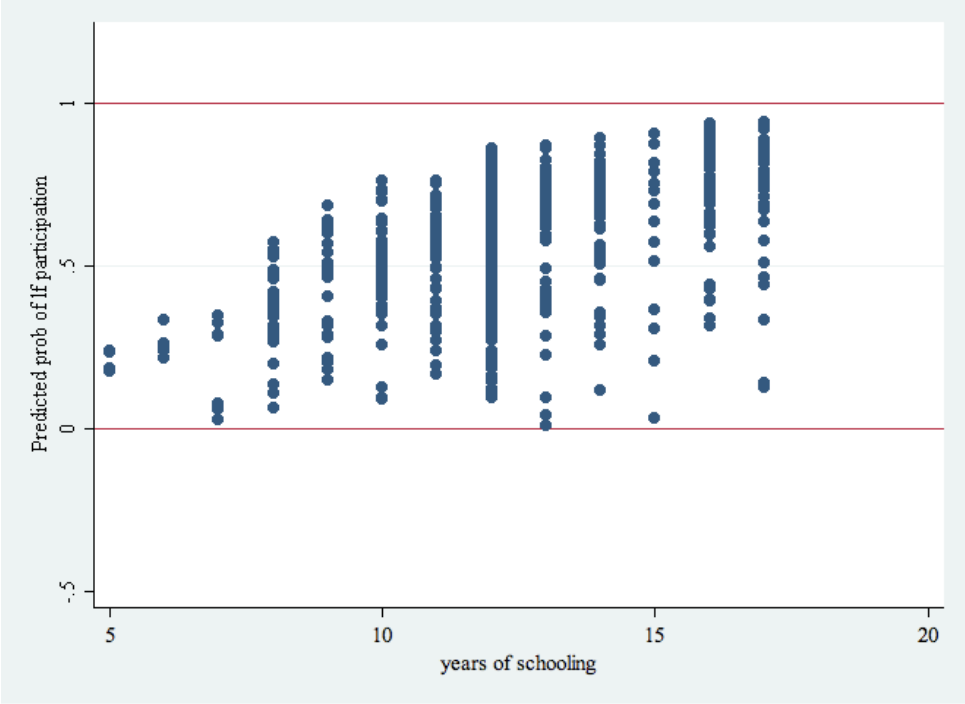
```
Obs:   753
inlf            byte   %9.0g                 =1 if in lab frce, 1975
age             byte   %9.0g                 woman's age in yrs
educ            byte   %9.0g                 years of schooling
kidslt6         byte   %9.0g                 # kids < 6 years
kidsge6         byte   %9.0g                 # kids 6-18
nwifeinc        float  %9.0g                 (faminc - wage*hours)/1000
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
city            byte   %9.0g                 =1 if live in SMSA
```

```
. use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ

. sum inlf age educ kidslt6 kidsge6 nwifeinc hushrs husage huseduc huswage unem city ,
sep(0)

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        inlf |       753    .5683931    .4956295          0          1
         age |       753    42.53785    8.072574         30         60
        educ |       753    12.28685    2.280246          5         17
     kidslt6 |       753    .2377158     .523959          0          3
     kidsge6 |       753    1.353254    1.319874          0          8
    nwifeinc |       753    20.12896     11.6348  -.0290575         96
      hushrs |       753    2267.271    595.5666        175       5010
      husage |       753    45.12085    8.058793         30         60
     huseduc |       753    12.49137    3.020804          3         17
     huswage |       753    7.482179    4.230559      .4121     40.509
        city |       753    .6427623    .4795042          0          1
```

2. Logit Model

```
. logit inlf nwifeinc educ age kidslt6 kidsge6 city
Iteration 0:   log likelihood =  -514.8732
Iteration 1:   log likelihood = -455.83688
Iteration 2:   log likelihood = -454.18791
Iteration 3:   log likelihood = -454.17931
Iteration 4:   log likelihood = -454.17931

Logistic regression                               Number of obs   =        753
                                                  LR chi2(6)      =     121.39
                                                  Prob > chi2     =     0.0000
Log likelihood = -454.17931                       Pseudo R2       =     0.1179

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0350754   .0080669    -4.35   0.000    -.0508862   -.0192646
        educ |   .2575602   .0409102     6.30   0.000     .1773777    .3377427
         age |  -.0576886   .0128004    -4.51   0.000    -.0827769   -.0326003
     kidslt6 |  -1.484777   .1980748    -7.50   0.000    -1.872996   -1.096558
     kidsge6 |  -.0666249   .0679011    -0.98   0.326    -.1997087    .0664589
        city |   .0191028     .17473     0.11   0.913    -.3233617    .3615672
       _cons |   .7254404   .7890909     0.92   0.358    -.8211493     2.27203
------------------------------------------------------------------------------
```

```
. mfx

Marginal effects after logit
      y  = Pr(inlf) (predict)
         =  .57425363
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.      z    P>|z|  [    95% C.I.   ]       X
---------+--------------------------------------------------------------------
nwifeinc |  -.0085755      .00197   -4.34   0.000  -.012446 -.004705    20.129
    educ |     .06297      .00999    6.30   0.000   .043387  .082553    12.2869
     age |  -.0141041      .00313   -4.51   0.000  -.020229 -.007979    42.5378
 kidslt6 |  -.3630078      .04862   -7.47   0.000  -.458302 -.267713    .237716
 kidsge6 |  -.0162889       .0166   -0.98   0.326  -.048825  .016248    1.35325
    city*|   .0046722      .04275    0.11   0.913  -.079123  .088467    .642762
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. predict inlfhat
(option p assumed; Pr(inlf))
. label variable inlfhat "Predicted prob of lf participation"
. twoway scatter inlfhat educ, yline(0 1) ylabel(-.5(.5)1.2)
```



Predicted values for the probability to be in the labor force

3. Test of multiple restrictions: The LR test
Do husband's characteristics influence the women's labor participation (beyond his income)?

```
. logit inlf nwifeinc educ age kidslt6 kidsge6 hushrs husage huseduc huswage city

Logistic regression                                     Number of obs   =        753
                                                        LR chi2(10)     =     129.27
                                                        Prob > chi2     =     0.0000
Log likelihood = -450.23675                             Pseudo R2       =     0.1255

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0182788   .0128726    -1.42   0.156    -.0435087    .0069511
        educ |   .2893468   .0478669     6.04   0.000     .1955294    .3831642
         age |  -.0383568   .0224972    -1.70   0.088    -.0824505    .0057369
     kidslt6 |  -1.537035    .200948    -7.65   0.000    -1.930886   -1.143184
     kidsge6 |  -.0648634   .0684488    -0.95   0.343    -.1990206    .0692938
      hushrs |  -.0003818   .0001706    -2.24   0.025    -.0007162   -.0000475
      husage |  -.0283468    .022439    -1.26   0.206    -.0723263    .0156328
     huseduc |  -.0354425   .0365281    -0.97   0.332    -.1070362    .0361511
     huswage |  -.0434876   .0372837    -1.17   0.243    -.1165623    .0295871
        city |   .0147352   .1809473     0.08   0.935     -.339915    .3693854
       _cons |   2.107232   .9407917     2.24   0.025     .2633139     3.95115
------------------------------------------------------------------------------
```

LR = 2 (loglikelihood UR – loglikelihood R) = 2 (-450.237+ 454.179)
chi2(4) = 7.89

Wald test:

```
. test hushrs husage huseduc huswage

 ( 1)  hushrs = 0
 ( 2)  husage = 0
 ( 3)  huseduc = 0
 ( 4)  huswage = 0

           chi2(  4) =    7.75
         Prob > chi2 =    0.1010
```

Presenting results

Table 1. Women labor force participation, 1975

	Mean	Base model (OLS)		Base model (LOGIT)		Extended models (LOGIT)			
	value	Coeff.	t-ratio	Marg. effect•	z*	Marg. effect•	z*	Marg. effect•	z*
Non-wife income ($1000)	20.1	-0.007	-4.52	-0.009	-4.42	-0.004	-1.42	-0.009	-4.34
Education (years)	12.3	0.053	6.72	0.063	6.34	0.071	6.05	0.063	6.32
Age (years)	42.5	-0.012	-4.59	-0.014	-4.53	-0.009	-1.71	-0.014	-4.45
Number of kids < 6 yrs old	0.24	-0.297	-8.29	-0.363	-7.47	-0.375	-7.62	-0.362	-7.45
Number of kids ≥ 6 yrs old	1.35	-0.012	-0.84	-0.016	-0.98	-0.016	-0.95	-0.016	-0.94
Husband hours of work ('000)	2.27					0.000	-2.25		
Husband's age	45.1					-0.007	-1.27		
Husband's education (years)	12.5					-0.009	-0.97		
Husband's hourly wage ($)	7.5					-0.011	-1.17		
District unmeployment (%)	8.6							-0.004	-0.68
Urban (1/0)	0.64							0.009	0.22
Intercept		0.645	4.01	-	-	-	-	-	-
Labor force participation (1/0)	0.568								
Number of observations	753								
R-squared / Chow R-squared		0.147		0.118		0.126		0.118	
Log(likelihood)				-454.2		-450.2		-453.9	

* z for the test of significativity of the underlying parameter
• Marginal effects computed at means of independent variables

- Number of observations
- Mean values of x and y (unless there is a table of descriptive stats), with units
- Coefficients / marginal effects + st. errors / t-ratio / p-value
- Overall goodness of fit: R^2, log-likelihood, pseudo-R^2