Department of Economics

Econometric Analysis of Panel Data

Professor William Greene
Phone: 212.998.0876
Office: KMC 7-78
Home page:www.stern.nyu.edu/~wgreene
Email: wgreene@stern.nyu.edu
URL for course web page:
www.stern.nyu.edu/~wgreene/Econometrics/PanelDataEconometrics.htm

Assignment 6
Nonlinear Models for Panel Data

Part I. A Concentrated Log Likelihood

Consider an exponential regression model with fixed effects, The density is

$f(y_{it}|\mathbf{x}_{it}) = [1/\theta_{it}]\exp(-y_{it} / \theta_{it}),\ y_{it} \geq 0$, where $\theta_{it} = \exp(\alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta})$, $i = 1,\ldots,n$; $t = 1,\ldots,T$.

It will prove convenient to let $\gamma_i = \exp(\alpha_i)$ so $\theta_{it} = \gamma_i \exp(\mathbf{x}_{it}'\boldsymbol{\beta}) = \gamma_i \lambda_{it}$.

The log likelihood for this exponential regression model with fixed effects is

(*) $$\log L = \sum_{i=1}^{n}\sum_{t=1}^{T}(-\log\theta_{it} - y_{it}/\theta_{it})$$

(a) Obtain the first order condition for maximizing log$L$ with respect to γ_i. Note, there is one of these for each i, so you need only differentiate

$$\log L_i = \sum_{t=1}^{T}(-\log\theta_{it} - y_{it}/\theta_{it})$$

with respect to γ_i and equate it to zero. You will gain some convenience by defining $a_{it} = y_{it}/\lambda_{it}$.

With the suggestion, the contribution of individual i to the log likelihood becomes

$$\log L_i = \sum_{t=1}^{T} (-\log\theta_{it} - y_{it}/\theta_{it})$$
$$= \sum_{t=1}^{T} (-\log\gamma_i - \log\lambda_{it} - (1/\gamma_i)(y_{it}/\lambda_{it})$$

The first derivative with respect to γ_i is

(**) $$\frac{\partial \log L_i}{\partial \gamma_i} = \sum_{t=1}^{T} \quad -\frac{1}{\gamma_i} + \frac{1}{\gamma_i^2} a_{it}$$

Equating this to zero, then multiplying both sides of the equation, we get

$$\gamma_i = \bar{a}_i$$

(This is the solution to b.)

(b) Now, treating $\boldsymbol{\beta}$ as if it were known, show that the implicit solution of this likelihood equation for γ_i in terms of $\boldsymbol{\beta}$ is

$$\gamma_i = \left(\frac{\Sigma_{t=1}^{T} \left[y_{it} / \exp(\mathbf{x}_{it}'\boldsymbol{\beta}) \right]}{T} \right) = \frac{\Sigma_{t=1}^{T} a_{it}}{T} = \bar{a}_i = \gamma_i(\boldsymbol{\beta})$$

See above

(c) It follows that at the solution for the MLE, it will be true that $\gamma_i(\beta) = \bar{a}_i$ where $\bar{a}_i$ is the sample mean of a_{it}. Denote $\theta_{it}^c = \bar{a}_i\lambda_{it}$. Insert this solution back into the log likelihood function, to obtain the *concentrated log likelihood function*

(***) $$\log L^c = \sum_{i=1}^{n} \sum_{t=1}^{T} (-\log\theta_{it}^c - y_{it}/\theta_{it}^c).$$

Note that this is a function of $\boldsymbol{\beta}$ but not of γ_i. To obtain the maximum likelihood estimator of $\boldsymbol{\beta}$, we can now maximize this function with respect to $\boldsymbol{\beta}$. This is equivalent to maximizing the whole log likelihood function, while considering only the solutions for γ_i that satisfy $\gamma_i = \gamma_i(\boldsymbol{\beta})$ as shown above. When we find β, we can then compute γ_i. (No assignment for this part.)

(d) With this in hand, it is now possible to maximize the function with respect to $\boldsymbol{\beta}$. Show that the likelihood equation will be

$$(****)\ \frac{\partial \log L^c}{\partial \beta} = \sum_{i=1}^{n}\sum_{t=1}^{T}\left(-\frac{1}{\theta_{it}^c}+\frac{y_{it}}{(\theta_{it}^c)^2}\right)\frac{\partial \theta_{it}^c}{\partial \beta} = \sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{\theta_{it}^c}\left(\frac{y_{it}}{\theta_{it}^c}-1\right)\frac{\partial \theta_{it}^c}{\partial \beta}.$$

You now need the derivative, $\partial\theta_{it}^c/\partial\boldsymbol{\beta}$. Continuing, show that

$\partial\theta_{it}^c/\partial\boldsymbol{\beta} = \bar{a}_i\lambda_{it}\mathbf{x}_{it} - \frac{1}{T}\sum_{t=1}^{T} a_{it}\lambda_{it}\mathbf{x}_{it}$ Hint: $\partial\lambda_{it}/\partial\beta = \lambda_{it}\mathbf{x}_{it}$. Insert your result in the log likelihood equation to obtain the implicit solution for $\boldsymbol{\beta}$,

$$\frac{\partial \log L^c}{\partial \beta} = \sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{\theta_{it}^c}\left(\frac{y_{it}}{\theta_{it}^c}-1\right)\left[\boldsymbol{\lambda_{it}}\bar{\mathbf{a}}_{\mathbf{i}}\mathbf{x}_{it} - \frac{1}{T}\Sigma_{t=1}^{T}\lambda_{it}a_{it}\mathbf{x}_{it}\right] = \mathbf{0}.$$

This involves a lot of tedious calculus. (***) results from simply inserting the solution for γ_i that was obtained in (a) into the original log likelihood function, (*). Because γ_i has been eliminated, this is the "concentrated" log likelihood. Likewise, the derivative in (****) just uses the chain rule in (**). The remaining complication is finding $\partial\theta_{it}^c/\partial\beta$, which is shown above.

Part II. Solving for FE in Panel Probit

For the binary fixed effects panel probit model,

$\text{Prob}(y_{it} = 1 \mid \mathbf{x}_{it}) = \Phi(\alpha_i + \mathbf{x}_{it}'\beta)$,
$\text{Prob}(y_{it} = 0 \mid \mathbf{x}_{it}) = 1 - \text{Prob}(y_{it} = 1 \mid \mathbf{x}_{it}) = \Phi(-\alpha_i - \mathbf{x}_{it}'\beta)$.

a. Write out the full log likelihood function.

The full log likelihood function is

$$\log L = \sum_{i=1}^{N}\left\{\left[\sum_{t,y_{it=0}}\log\Phi(-\alpha_i - \mathbf{x}_{it}'\beta)\right] + \left[\sum_{t,y_{it=1}}\log\Phi(\alpha_i + \mathbf{x}_{it}'\beta)\right]\right\}$$

A very useful simplification is to write $q_{it} = 2y_{it}-1$, so that $q_{it} = 1$ when $y_{it} = 1$ and -1 when $y_{it} = 0$. Then,

$$\log L = \sum_{i=1}^{N}\sum_{t=1}^{T}\log\Phi\left[q_{it}(\alpha_i + \mathbf{x}_{it}'\beta)\right]$$

b. Write out the first order condition for maximizng the function with respect to α_i, taking $\boldsymbol{\beta}$ as known. Take this derivation as far as possible – you will ultimately find that unlike the exponential model we examined in class, in this model, there is no explicit solution for α_i in terms of $\boldsymbol{\beta}$ and the data.

In the expression above, only T terms in the inner sum involve α_i, That part of the log likelihood function is

$$\log L_i = \sum_{t=1}^{T} \log \Phi\left[q_{it}(\alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta})\right]$$

$$\frac{\log L_i}{\partial \alpha_i} = \sum_{t=1}^{T} \frac{1}{\Phi\left[q_{it}(\alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta})\right]} \phi\left[q_{it}(\alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta})\right] q_{it}$$

No further simplification is possible.

c. Show that regardless of the finding in b, there is no solution for α_i when y_{it} is always 1 or always 0 within a given group (i).

In the derivative, both ϕ and Φ are always positive. The terms are negative when $y_{it} = 0$ and positive when $y_{it} = 1$. So, if y_{it} is always equal to 1, then the sum has to be positive and you cannot equate it to 0. If y_{it} is always 0, then the sum has to be negative and, again, you cannot equate it to zero.

Part III. The Incidental Parameters Problem.

This is a purely empirical exercise. It will involve some computations using the German health care data.

As we discussed in class, for the binary logit model, there are two estimators for the fixed effects model

$$\text{Prob}(y_{it} = 1 \mid x_{it}) = \Lambda(\alpha_i + \beta' x_{it}),\ i = 1,...,n,\ t = 1,...,T.$$

The ‘brute force’ approach maximizes the whole log likelihood for α_i,i = 1,...,n and β. This estimator is known to suffer from the ‘incidental parameters problem;’ when T is small, the estimator is biased away from zero. The best known result is that when T = 2, there is a 100% bias. The other approach is the Rasch/Andersen/Chamberlain method, which computes a conditional MLE using the probabilities conditioned on the sum of the y_{it}s for each group. This estimator is known to be consistent. For this exercise, we will see if the effect is visible in a sample, using precisely the estimators described.

a. We first see if we can observe Hsiao/Abrevaya's finding when T = 2. The following commands compute the estimates of the logit model both ways. Estimate the equations, and report your results. Do the empirical results seem to conform to the theory?

```
Sample ; All $
Reject ; _groupti # 2 $
Namelist ; x = hhninc,age,married,working$
Create ; y = doctor $
Logit ; Lhs = y ; Rhs = x ; pds = 2 ; Fixed$
Logit ; Lhs = y ; Rhs = x ; pds = 2 $
POOLED
------------------------------------------------------------------
Logit    Regression Start Values for Y
Dependent variable                    Y
Log likelihood function     -1378.51484
Estimation based on N =   2158, K =   5
Information Criteria: Normalization=1/N
--------+---------------------------------------------------------
        |                  Standard            Prob.       Mean
       Y| Coefficient        Error       z    z>|Z|       of X
--------+---------------------------------------------------------
  HHNINC|    -.29336         .26910    -1.09  .2756      .35678
     AGE|     .02488***      .00403     6.18  .0000    38.9222
 MARRIED|    -.03152         .09869     -.32  .7494      .61538
 WORKING|    -.10627         .10416    -1.02  .3076      .62326
Constant|    -.17427         .19791     -.88  .3786
--------+---------------------------------------------------------
BRUTE FORCE
------------------------------------------------------------------
FIXED EFFECTS Logit  Model
Dependent variable                    Y
Log likelihood function      -474.89737
Estimation based on N =   2158, K = 349
Sample is  2 pds and   1079 individuals
Skipped  734 groups with inestimable ai   <========================
LOGIT (Logistic) probability model
--------+---------------------------------------------------------
        |                  Standard            Prob.       Mean
       Y| Coefficient        Error       z    z>|Z|       of X
--------+---------------------------------------------------------
        |Index function for probability
  HHNINC|   -1.79046        1.14804    -1.56  .1189      .35928
     AGE|     .00177         .05837      .03  .9758    36.2652
 MARRIED|     .52879         .45144     1.17  .2415      .56667
 WORKING|     .71611*        .37067     1.93  .0534      .69275
--------+---------------------------------------------------------
CONDITIONAL
------------------------------------------------------------------
Logit Model for Panel Data
Dependent variable                    Y
Log likelihood function      -237.44868
Estimation based on N =   2158, K =   4
Fixed Effect Logit Model for Panel Data
--------+---------------------------------------------------------
        |                  Standard            Prob.
       Y| Coefficient        Error       z    z>|Z|
--------+---------------------------------------------------------
  HHNINC|    -.89523         .81179    -1.10  .2701
     AGE|     .00089         .04127      .02  .9829
 MARRIED|     .26440         .31921      .83  .4075
 WORKING|     .35806         .26210     1.37  .1719
--------+---------------------------------------------------------
```

Three sets of results are given. The pooled results look quite far off compared to the third set of results which we know are from a consistent estimator. Comparing the second set of results to the third, it looks like the theory is working very well. The unconditional estimates are almost exactly twice the conditional estimates.

b. A second result that seems intuitively reasonable is that the IP bias diminishes as T increases. Is this the case? Change the three 2s in the command set above to 3s and redo the experiment. What do you find? Now, change the 2s to 7s and repeat the experiment. In each case, report your findings and your conclusions.

Here is the full experiment with T=2,3,4,5,6,7

```
WITH T=2
--------+-------------------------------------------------------------
        |                   Standard            Prob.        Mean
       Y| Coefficient        Error       z    z>|Z|         of X
--------+-------------------------------------------------------------
        |Index function for probability
  HHNINC|   -1.79046        1.14804    -1.56  .1189      .35928
     AGE|     .00177         .05837      .03  .9758     36.2652
 MARRIED|     .52879         .45144     1.17  .2415      .56667
 WORKING|     .71611*        .37067     1.93  .0534      .69275
--------+-------------------------------------------------------------
  HHNINC|    -.89523         .81179    -1.10  .2701
     AGE|     .00089         .04127      .02  .9829
 MARRIED|     .26440         .31921      .83  .4075
 WORKING|     .35806         .26210     1.37  .1719
--------+-------------------------------------------------------------
WITH T=3
--------+-------------------------------------------------------------
        |                   Standard            Prob.        Mean
       Y| Coefficient        Error       z    z>|Z|         of X
--------+-------------------------------------------------------------
        |Index function for probability
  HHNINC|     .59013         .70428      .84  .4021      .36352
     AGE|     .12338***      .03485     3.54  .0004     40.4401
 MARRIED|     .65030*        .35210     1.85  .0648      .74479
 WORKING|     .09112         .27020      .34  .7359      .66146
--------+-------------------------------------------------------------
  HHNINC|     .40128         .57716      .70  .4869
     AGE|     .08214***      .02841     2.89  .0038
 MARRIED|     .42203         .28313     1.49  .1361
 WORKING|     .05916         .22023      .27  .7882
--------+-------------------------------------------------------------
WITH T=4
--------+-------------------------------------------------------------
        |                   Standard            Prob.        Mean
       Y| Coefficient        Error       z    z>|Z|         of X
--------+-------------------------------------------------------------
        |Index function for probability
  HHNINC|   -1.27207**       .53695    -2.37  .0178      .36115
     AGE|     .17885***      .02282     7.84  .0000     41.6709
 MARRIED|    -.32722         .25114    -1.30  .1926      .73537
 WORKING|     .17151         .19681      .87  .3835      .68236
--------+-------------------------------------------------------------
  HHNINC|    -.94006**       .45951    -2.05  .0408
     AGE|     .13319***      .01955     6.81  .0000
 MARRIED|    -.24378         .21610    -1.13  .2593
 WORKING|     .12651         .16941      .75  .4552
--------+-------------------------------------------------------------
WITH T=5
--------+-------------------------------------------------------------
        |                   Standard            Prob.        Mean
       Y| Coefficient        Error       z    z>|Z|         of X
--------+-------------------------------------------------------------
        |Index function for probability
  HHNINC|    -.38665         .38487    -1.00  .3151      .35799
     AGE|     .11721***      .01661     7.06  .0000     42.9664
 MARRIED|    -.17447         .22345     -.78  .4349      .78049
 WORKING|     .18235         .16158     1.13  .2591      .72052
--------+-------------------------------------------------------------
  HHNINC|    -.30921         .34404     -.90  .3688
     AGE|     .09371***      .01483     6.32  .0000
 MARRIED|    -.14012         .19971     -.70  .4829
 WORKING|     .14476         .14422     1.00  .3155
--------+-------------------------------------------------------------
WITH T=6
```

Y	Coefficient	Standard Error	z	Prob. z>|Z|	Mean of X
	Index function for probability				
HHNINC	.76162**	.35975	2.12	.0343	.35924
AGE	.08686***	.01337	6.49	.0000	44.1876
MARRIED	-.44181*	.22544	-1.96	.0500	.80129
WORKING	-.52941***	.14991	-3.53	.0004	.74873
HHNINC	.63224*	.32811	1.93	.0540	
AGE	.07233***	.01220	5.93	.0000	
MARRIED	-.36869*	.20575	-1.79	.0731	
WORKING	-.44155***	.13683	-3.23	.0013	

WITH T=7

Y	Coefficient	Standard Error	z	Prob. z>|Z|	Mean of X
	Index function for probability				
HHNINC	-.03468	.33461	-.10	.9175	.34717
AGE	.08870***	.01193	7.43	.0000	43.8163
MARRIED	-.04349	.21200	-.21	.8374	.84234
WORKING	-.09784	.14573	-.67	.5020	.76362
HHNINC	-.03061	.30848	-.10	.9210	
AGE	.07604***	.01103	6.89	.0000	
MARRIED	-.03777	.19604	-.19	.8472	
WORKING	-.08394	.13482	-.62	.5336	

The estimators are clearly getting closer together as T ncreases. With T=7, they seem to be pretty close. The general experience suggests that if T is 10 or larger, the difference is small enough to be comfortable with the unconditional estimator. One might think that this is not necessarily a useful result, since we can always use the conditional estimator. But, the conditional estimator only exsts for the logit model. If you want to use a probit model, you can only use the unconditional estimator, so this is a useful guide for that case.

c. What do you conclude about the fixed effects model?

See above.

Part IV. A Common Effects Probit Model

In this exercise, you will fit a probit model with common effects, and develop the appropriate model based on your findings. The probit model we will use is

$\text{Prob}(y_{it} = 1 \mid x_{it}) = \Phi(c_i + \beta' x_{it})$
y_{it} = Public_{it} = whether or not the individual chose public health insurance in that year.
x_{it} = one,age,educ,hhninc,handper,working,hsat

1. Suppose, for the moment, we ignore the heterogeneity, c_i and just pool the data and fit a simple probit model. Is the estimator consistent? What assumptions are necessary for the pooled estimator to be a consistent estimator of β?

The only way for the pooled estimator to be a consistent estimator of β is for c_i to equal zero for every observation. Otherwise, it is inconsistent in all cases.

2. All of the suggested covariates in the model are time varying. Fit a random effects model and a fixed effects model (this can only be done by brute force). Report your results.

The pooled, random and fixed effects results are shown below.

```
--> probit;lhs=public;rhs=xit;random;pds=_groupti$
-----------------------------------------------------------------
Binomial Probit Model
Dependent variable               PUBLIC
Log likelihood function     -8294.31338
Restricted log likelihood   -9711.25153
Chi squared [   6 d.f.]      2833.87629
Significance level               .00000
McFadden Pseudo R-squared      .1459069
Estimation based on N =  27326, K =   7
--------+--------------------------------------------------------
        |                  Standard            Prob.       Mean
  PUBLIC| Coefficient        Error       z    z>|Z|       of X
--------+--------------------------------------------------------
        |Index function for probability
Constant|    3.86731***      .08308    46.55  .0000
     AGE|    -.00034         .00106     -.32  .7498    43.5257
    EDUC|    -.16849***      .00407   -41.37  .0000    11.3206
  HHNINC|    -.96505***      .05584   -17.28  .0000     .35208
 HANDPER|     .00114         .00070     1.64  .1014    7.01229
 WORKING|    -.01036         .02533     -.41  .6825     .67705
    HSAT|    -.03834***      .00532    -7.21  .0000    6.78543
--------+--------------------------------------------------------
Normal exit:  24 iterations. Status=0. F=    4868.491

-----------------------------------------------------------------
Random Effects Binary Probit Model
Dependent variable               PUBLIC
Log likelihood function     -4868.49090
Restricted log likelihood   -8294.31338
Chi squared [   1 d.f.]      6851.64496
Significance level               .00000
Unbalanced panel has   7293 individuals
--------+--------------------------------------------------------
        |                  Standard            Prob.       Mean
  PUBLIC| Coefficient        Error       z    z>|Z|       of X
--------+--------------------------------------------------------
Constant|    11.5697***      .33006    35.05  .0000
     AGE|    -.00291         .00362     -.80  .4214    43.5257
    EDUC|    -.56234***      .01755   -32.04  .0000    11.3206
  HHNINC|   -1.46850***      .13213   -11.11  .0000     .35208
 HANDPER|     .00254         .00205     1.24  .2157    7.01229
 WORKING|     .10990*        .06136     1.79  .0733     .67705
```

```
    HSAT|    -.03535**       .01383    -2.56  .0106    6.78543
     Rho|     .90503***      .00416   217.77  .0000
--------+-----------------------------------------------------------
```

--> probit;lhs=public;rhs=xit;fem;pds=_groupti$

```
--------------------------------------------------------------------
FIXED EFFECTS Probit Model
Dependent variable               PUBLIC
Log likelihood function     -1346.22838
Estimation based on N =  27326, K =1236
Information Criteria: Normalization=1/N
              Normalized   Unnormalized
AIC               .18899     5164.45676
Fin.Smpl.AIC      .19328     5281.66570
Bayes IC          .56060    15318.93084
Hannan Quinn      .30876     8437.17555
Model estimated: Feb 10, 2010, 07:04:22
Unbalanced panel has   7293 individuals
Skipped 6063 groups with inestimable ai
PROBIT (normal)  probability model
--------+-----------------------------------------------------------
        |                  Standard            Prob.       Mean
  PUBLIC| Coefficient        Error       z    z>|Z|       of X
--------+-----------------------------------------------------------
        |Index function for probability
     AGE|    -.04549***      .01178    -3.86  .0001    42.1510
    EDUC|    -.30999***      .08493    -3.65  .0003    12.5433
  HHNINC|    -.94060***      .25739    -3.65  .0003     .39798
 HANDPER|     .00017         .00437      .04  .9692    5.53937
 WORKING|     .09822         .11254      .87  .3828     .78911
    HSAT|    -.02955         .02123    -1.39  .1639    7.17090
--------+-----------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
--------------------------------------------------------------------
```

3. We are interested in deciding which is preferred, fixed or random effects. I propose to use a variable addition test. Add the group means to the model, then carry out a likelihood ratio test of the hypothesis that the coefficients on the group means are all zero. What do you find? What do you conclude is the preferred model?

```
sample;all$
namelist;xit =  one,age,educ,hhninc,handper,working,hsat$
probit;lhs=public;rhs=xit;fem;pds=_groupti$
calc;logl0=logl$
create ; agebar=GroupMean(age,pds=_groupti) $
create ; edbar=GroupMean(educ,pds=_groupti) $
create ; incbar=GroupMean(hhninc,pds=_groupti) $
create ; handbar=GroupMean(handper,pds=_groupti) $
create ; workbar=GroupMean(working,pds=_groupti) $
create ; hsatbar=GroupMean(hsat,pds=_groupti) $
namelist;xb=agebar,edbar,incbar,handbar,workbar,hsatbar$
probit ; lhs=public;rhs=xit,xb;random;pds=_groupti $
calc;logl1=logl$
calc;list;chisq=2*(logl1-logl0);ctb(.95,6)$
matrix;bm=b(8:13);vm=varb(8:13,8:13)
;list;waldstat=bm'<vm>bm$
```

Here are the results for the model with the group means added. The log likelihood without the group means is -4868.49090. With the group means, it is -4830.45663. Twice the difference is 76.06854. This is a chi squared statistic with 6 degrees of freedom. The 95% critical value is 12.5915, so the hypothesis that the coefficients on the group means are all zero would be rejected. This makes the fixed effects model the preferred specification. The Wald statistic is 172.82994 which leads to the same inference.

```
-----------------------------------------------------------------
Random Effects Binary Probit Model
Dependent variable               PUBLIC
Log likelihood function     -4830.45663
Restricted log likelihood   -8231.41932
Chi squared [   1 d.f.]      6801.92539
Significance level               .00000
McFadden Pseudo R-squared      .4131684
Estimation based on N =  27326, K =  14
Information Criteria: Normalization=1/N
              Normalized   Unnormalized
AIC               .35457     9688.91326
Fin.Smpl.AIC      .35457     9688.92863
Bayes IC          .35878     9803.93157
Hannan Quinn      .35592     9725.98289
Model estimated: Feb 10, 2010, 07:14:13
Unbalanced panel has   7293 individuals
--------+--------------------------------------------------------
        |                  Standard            Prob.       Mean
  PUBLIC| Coefficient        Error       z    z>|Z|       of X
--------+--------------------------------------------------------
Constant|    12.7702***      .42367    30.14  .0000
     AGE|    -.03087***      .00836    -3.69  .0002    43.5257
    EDUC|    -.24376***      .05281    -4.62  .0000    11.3206
  HHNINC|    -.67317***      .19869    -3.39  .0007     .35208
 HANDPER|     .00037         .00356      .10  .9167    7.01229
 WORKING|     .05389         .08190      .66  .5105     .67705
    HSAT|    -.01904         .01723    -1.10  .2693    6.78543
  AGEBAR|     .03372***      .00888     3.80  .0001    43.5257
   EDBAR|    -.26090***      .05530    -4.72  .0000    11.3206
  INCBAR|   -3.26485***      .33218    -9.83  .0000     .35208
 HANDBAR|     .00519         .00453     1.15  .2517    7.01229
 WORKBAR|    -.00795         .13099     -.06  .9516     .67705
 HSATBAR|    -.17811***      .03074    -5.79  .0000    6.78543
     Rho|     .90640***      .00445   203.59  .0000
--------+--------------------------------------------------------

Matrix WALDSTAT has  1 rows and  1 columns.
               1
        +-------------+
       1|  172.82994
        +-------------+
```

4. Suppose it were hypothesized that the previous year's choice of whether or not to choose public insurance were on the right hand side of the equation. That is,

$$\text{Prob}(y_{it} = 1 \mid x_{it}) = \Phi(c_i + \beta' x_{it} + \gamma y_{i,t-1})$$

What would this imply for how one (you) should go about estimating the parameters of the model. What issues should you be concerned with for a dynamic model?

The new specification creates a problem for estimation. Even if it is argued that c_i is uncorrelated with x_{it}, it can't be uncorrelated with y_{it-1}. So, some alternative approach would be called for. Wooldridge's suggestion is a combination of Mundlak's approach and a separate model for the initial conditions, y_{i0}. In one way or another, the model will have to account for the endogeneity of the lagged dependent variable in the equation.