THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

ECON 7320 Advanced Microecnometrics

Assignment II

Student name: Chi-Sheng Lo
Student ID number: 44046918

Q1. MLE/LOGIT

Given y = $\begin{cases} 1 \; with \; probability \; \Lambda(\beta_0 + \beta_1 x) \\ 0 \; with \; probability \; 1 - \Lambda(\beta_0 + \beta_1 x) \end{cases}$

Where $\Lambda()$ is the logistic cdf.

	y = 0	y = 1
x = 0	52	85
x = 1	51	12

(i) Obtain the Maximum Likelihood Estimator of $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$

Likelihood:

L = $\prod_{i=1}^{n} p(x)^y \, (1 - p(x))^{1-y}$ = $\prod_{i=1}^{n} \Lambda(\beta_0 + \beta_1 x)^y \, (1 - \Lambda(\beta_0 + \beta_1 x))^{1-y}$

Log-likelihood turns products into sums:

$\ell = \sum_{i=1}^{n} ylog$p(x) + (1 - y)log1$-$ p(x)

$= \sum_{i=1}^{n} log 1 - p(x) + \sum_{i=1}^{n} ylog \frac{p(x)}{1-p(x)}$

$= \sum_{i=1}^{n} log 1 - p(x) + \sum_{i=1}^{n} y(\beta_0 + x\beta_1)$

$= \sum_{i=1}^{n} - log 1 \; + \; e^{\beta_0 + x\beta_1} + \sum_{i=1}^{n} y(\beta_0 + x\beta_1)$

Where:

P(x) = $(1 + \; e^{-x_i^T \beta})^{-1}$ and 1 – P (x) = $\frac{e^{-x_i^T \beta}}{(1+ e^{-x_i^T \beta})}$

Then:

ln P(x) = – ln(1 + $e^{-x_i^T \beta}$) and ln(1-P(x)) = – $x_i^T \beta$ – ln(1 + $e^{-x_i^T \beta}$)

The log-likelihood function is rewritten as:

$\ell = \sum_{i=1}^{n}[(y_i - 1) \; x_i^T \beta$ – ln(1 + $e^{-x_i^T \beta}$)]

Taking the gradient of the log-likelihood function, we obtain the score function:

S(β; y) = $\nabla_\beta \; \ell(\beta$; y) = $\sum_{i=1}^{n}[(y_i - 1)x_i + \frac{x_i e^{-x_i^T \beta}}{(1+ e^{-x_i^T \beta})}]$

$= \sum_{i=1}^{n}[(y_i - (1 + \; e^{-x_i^T \beta})^{-1}]x_i$

$= \sum_{i=1}^{n}(y_i -$ P(x))x_i

Differentiating the score function with respect to β and multiplying by -1, we obtain the observed information matrix:

$$\mathrm{I}(\beta;\, \mathrm{y}) = -\nabla_\beta^2(\beta;\, \mathrm{y}) = \sum_{i=1}^{n} \frac{x_i e^{-x_i^T \beta}}{(1+e^{-x_i^T \beta})} x_i x_i^T = \sum_{i=1}^{n} \mathrm{P}(x)(1-\mathrm{P}(x))\, x_i x_i^T$$

More specifically:

$$\mathrm{I}(\beta;\, \mathrm{y}) = -\begin{bmatrix} \frac{\partial^2 \ell(\beta_0,\beta_1)}{\partial \beta_0 \partial \beta_0^T} & \frac{\partial^2 \ell(\beta_0,\beta_1)}{\partial \beta_0 \partial \beta_1^T} \\ \frac{\partial^2 \ell(\beta_0,\beta_1)}{\partial \beta_1^T \beta_0} & \frac{\partial^2 \ell(\beta_0,\beta_1)}{\partial \beta_1 \partial \beta_1^T} \end{bmatrix}$$

Thereafter, the iteration of Fisher's scoring method can be shown as:

$\beta_t = \beta_{t-1} + [I(\beta_{t-1})]^{-1}\, \mathrm{S}(\beta_{t-1}; y)$ will converge to β^* from any initial point β_0 in that neighborhood. After the MLE $\hat{\beta}$ is obtained, we can then compute the asymptotic covariance matrix as $I^{-1}(\hat{\beta})$.

As a result:

$\widehat{\beta_0}$ = 0.491

$\widehat{\beta_1}$ = – 1.938

(ii) Obtain the estimated asymptotic standard errors

If x is continuous, then

$\frac{\partial p(x)}{\partial x}$ = g(xβ)β_j

Where:

g(z) = $\frac{dG}{dz}$ (z)

G(z) = Λ(z) = exp(z)/[1 + exp(z)]

P(y =1|x) = G(xβ) = p(x)

Thereafter, we should use the score of the conditional log likelihood:

$$s_i(\beta) = \frac{g(x_i\beta) x_i^T [y_i - G(x_i\beta)]}{G(x_i\beta)[1 - G(x_i\beta)]}$$

Then the expected value of the Hessian conditional on x_i:

$$-\mathrm{E}[\mathrm{H}_i(\beta) |\, x_i] = \frac{[g(x_i\beta)]^2 x_i^T x_i}{G(x_i\beta)[1 - G(x_i\beta)]} = \mathrm{A}(x_i,\, \beta)$$

Av$\hat{a}$r($\hat{\beta}$) = $\{\sum_{i=1}^{n} \frac{[g(x_i\widehat{\beta})]^2 x_i^T x_i}{G(x_i\widehat{\beta})[1-\widehat{G(x_i\beta)}]}\}^{-1}$ = $\hat{V}$

Where: $\hat{V}$ is positive definite

Taking the square root of the jth diagonal element of $\hat{V}$ should we get to the asymptotic standard error of $\widehat{\beta_1}$ which is 0.3660 and asymptotic standard error of $\widehat{\beta_0}$ which is 0.1761.

(iii) Test the hypothesis that $\boldsymbol{\beta_1}$ = 0 using a Wald test

Null hypothesis:

H_0: β_1 = 0

Alternative hypothesis:

H_1: $\beta_1 \neq 0$

Wald statistics:

W = c($\hat{\beta})^T[C(\hat{\beta})(\widehat{\sigma^2}(x^T x)^{-1})C(\hat{\beta})^T]^{-1}c(\hat{\beta}) \xrightarrow{d} \chi^2$[J].

Where c($\hat{\beta}$) = $\frac{\partial c(\hat{\beta})}{\partial \hat{\beta}^T}$

We reject the H_0 if the p-value p = Pr{χ^2[J] > W} < 0.05.

Conclusion from the Wald test:

Since W = $(-\ 1.938)^2 \times ((0.366)^2)^{-1}$ = 28.05 which is larger than the two-sided critical values of α = 0.025 at 5.02, we can reject the H_0 that β_1 = 0 and conclude that $\beta_1 \neq 0$ at five percent significance level.

(iv) Test the hypothesis that $\boldsymbol{\beta_1}$ = 0 using a likelihood ratio test

Null hypothesis:

H_0: β_1 = 0

Alternative hypothesis:

H_1: $\beta_1 \neq 0$

The test statistics and its limiting distribution under H_0:

Likelihood ratio (LR) test = – 2[ln $L_* -$ ln L] $\xrightarrow{d} \chi^2$[J].

We reject the H_0 if the p-value p = Pr{χ^2[J] > LR} < 0.05.

Conclusion from the LR test:

Since LR = 33.83 which is larger than the two-sided critical values of α = 0.025 at 5.02, we can reject the H_0 that β_1 = 0 and conclude that $\beta_1 \neq 0$ at five percent significance level.

(v) Compute the marginal effect of $\boldsymbol{\beta_1}$ evaluated at x = 1

Probability Pr = $\Lambda(x'\,\beta) = \frac{e^{x'B}}{(1+e^{x'B})}$

Marginal effect = $\frac{\partial Pr(y=1)}{\partial x_1} = \frac{e^{x\beta}}{(1+e^{x'B})^2}\frac{\partial(x\beta)}{\partial x_1}$

$= \frac{e^{x\beta}}{(1+e^{x'B})^2}\,\beta_1$

$= \Lambda(x'\beta)(1 - \Lambda(x'\beta))\,\beta_1$

= Pr(y = 1|x = 1) Pr(y = 0|x = 1) β_1

$= \frac{51}{63} \times \frac{12}{63} \times -1.938 = -0.299$

Verification by STATA:

```
. logit y x

Iteration 0:   log likelihood = -138.53942
Iteration 1:   log likelihood = -121.69588
Iteration 2:   log likelihood = -121.62287
Iteration 3:   log likelihood = -121.62286

Logistic regression                               Number of obs     =        200
                                                  LR chi2(1)        =      33.83
                                                  Prob > chi2       =     0.0000
Log likelihood = -121.62286                       Pseudo R2         =     0.1221

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |  -1.938326   .3659735    -5.30   0.000    -2.655621   -1.221031
       _cons |   .4914075   .1760553     2.79   0.005     .1463454    .8364696
------------------------------------------------------------------------------
```

```
. //marginal effect at x=1//
. quietly logit y x

. margins, dydx(*) at (x=1)

Conditional marginal effects                    Number of obs     =        200
Model VCE    : OIM

Expression   : Pr(y), predict()
dy/dx w.r.t. : x
at           : x               =           1

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |  -.2988803   .0288924   -10.34   0.000    -.3555084   -.2422522
------------------------------------------------------------------------------
```

Q2. Structural Models

(i) Show that $\mathbf{I_2}$ – A is an invertible matrix

$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = (1 \times 1) - (0 \times 0) = 1$

$A = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix} = (0 \times 0) - (\alpha \times 0) = 0$

$I_2 - A = 1 - 0 = 1$

Since $I_2 - A \neq 0$ which is non-singular. This is the key assumption that proves that $I_2 - A$ is invertible.

(ii) Show that Y = $(\mathbf{I_2} - \mathbf{A})^{-1}\boldsymbol{\varepsilon}$

Again, since: $I_2 - A = 1 - 0 = 1$ (Invertible)

$Y = (I_2 - A)^{-1}\varepsilon = (I_2 - A)\ \varepsilon = \varepsilon$

$\varepsilon = Y(I_2 - A) = Y - YA$

$\varepsilon = Y\ (1 - A)$

$Y = \varepsilon(1 - A)^{-1} = (I_2 - A)^{-1}\varepsilon$

(iii) Show that E[Y] = 0 and VAR[Y] = Π where $\boldsymbol{\Pi = \sigma^2(I_2 - A)^{-1}(I_2 - A^T)^{-1}}$

Since $E[\varepsilon] = 0$,

$E[Y] = E[(I_2 - A)^{-1}\varepsilon] = E[(I_2 - A)^{-1}]\ E[\varepsilon] = E[(I_2 - A)^{-1}] \times 0 = 0$

$VAR[Y] = [(I_2 - A)^{-1}\varepsilon] = VAR[\varepsilon]\ (I_2 - A)^{-1}((I_2 - A)^{-1})^T$

$= \sigma^2(I_2 - A)^{-1}(I_2 - A^T)^{-1} = \sigma^2 \times 1 \times 1 = \Pi$

(iv) Show that $\Pi = \begin{pmatrix} \boldsymbol{\sigma^2(1+\ \sigma^2)} & \boldsymbol{\alpha\sigma^2} \\ \boldsymbol{\alpha\sigma^2} & \boldsymbol{\sigma^2} \end{pmatrix}$. Are the structural parameters α and $\boldsymbol{\sigma^2}$ identified?

Recaptures: $(I_2 - A)^{-1} = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ and $(I_2 - A^T)^{-1} = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}$

Thereafter:

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} = \begin{pmatrix} 1 \times 1 + \alpha \times \alpha & 1 \times 0 + \alpha \times 1 \\ 0 \times 1 + 1 \times \alpha & 0 \times 0 + 1 \times 1 \end{pmatrix} = \begin{pmatrix} 1 + \alpha^2 & \alpha \\ \alpha & 1 \end{pmatrix}$$

$$\Pi = \sigma^2 \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 + \alpha^2 & \alpha \\ \alpha & 1 \end{pmatrix} = \begin{pmatrix} \sigma^2 + \sigma^2\alpha^2 & \alpha\sigma^2 \\ \alpha\sigma^2 & \sigma^2 \end{pmatrix}$$

Yes, the structural parameters are identified from $\sigma^2 = \Pi_{22}$ and $\alpha = \frac{\Pi_{12}}{\Pi_{22}} = \frac{\alpha\sigma^2}{\sigma^2}$

(v) Suppose now that we have an independent random sample $(\mathbf{Y_i})_{\boldsymbol{i=1}}^{\boldsymbol{N}}$. Propose a consistent estimator of Π

Given E[Y] = 0 and Y = $(1\ -\ 0)^{-1}$ ε = $(I_2 - A)$ ε.

Therefore: VAR[Y] = E[$(I_2 - A)$]E[ε $(I_2 - A)\varepsilon^T$] = E[YY^T]

For Π to be consistent: It is $\widehat{\Pi} = \frac{1}{N}\sum_{i=1}^{N} Y_i\, Y_i^T$ which means that Π is consistent as N gets larger. Thus, plim $\widehat{\Pi}$ = Π.

(vi) Use your consistent estimator of Π to propose consistent estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma^2}$.

Since the consistent estimator of Π is $\widehat{\Pi} = \frac{1}{N}\sum_{i=1}^{N} Y_i\, Y_i^T$, then the consistent estimator of α is $\hat{\alpha} = \frac{1}{N}\sum_{i=1}^{N} \frac{\alpha_i\sigma_i^2}{\sigma_i^2} = \frac{\widehat{\Pi_{12}}}{\widehat{\Pi_{22}}}$ and the consistent estimator of σ^2 is $\widehat{\sigma^2} = \frac{1}{N}\sum_{i=1}^{N} \sigma^2 = \widehat{\Pi_{22}}$. Therefore, α and σ^2 can be consistent as N gets larger.

Q3. Dynamic Panel Data Model

Given that:

y_{it} = βy_{it-1} + u_{i1} , I = 1 ,…, N, t = 1, …, T (1)

Where u_{it} = α_i + ε_{it}

E[ε_{it}] = 0, E[ε_{it}]^2 = σ_ε^2 , E[$\varepsilon_{it}\varepsilon_{is}$] = 0 for t ≠ s, E[$\alpha_i\varepsilon_{it}$] = 0,

E[α_i^2] = σ_α^2 and | β| < 1 and T ≥ 3.

(i) Write down y_{i1} assuming that y_{i0} = 0.

y_{i1} = βy_{i0} + u_{i1}

Given that:

y_{i0} = 0

u_{i1} = α_i + ε_{i1}

Then:

y_{i1} = 0 + u_{i1} = α_i + ε_{i1}

(ii) Compute y_{i2} as a function of α_i, ε_{i1} and ε_{i2}

y_{i2} = βy_{i1} + u_{i2} = β (βy_{i0} + α_i + ε_{i1}) + (α_i + ε_{i2})

If y_{i0} = 0, then:

y_{i2} = βy_{i1} + u_{i2} = β(α_i + ε_{i1}) + (α_i + ε_{i2})

(iii) Show that y_{it} = $\alpha_i\left(\frac{1-\beta^t}{1-\beta}\right)$ + $\sum_{s=1}^{t}\beta^{t-s}\,\varepsilon_{is}$, i=1,…,N, t=1,…,T

y_{it} = $\beta y_{i,t-1}$ + α_i + ε_{it}

= $\beta^2 y_{i,t-2}$ + α_i(1+ β) + ε_{it} + + $\beta\varepsilon_{i,t-1}$

= $\beta^3 y_{i,t-3}$ + α_i(1+ $\beta + \beta^2$) + ε_{it} + + $\beta\varepsilon_{i,t-1}$ + $\beta^2\varepsilon_{i,t-2}$

= …

= $\beta^t y_{i0}$ + $\alpha_i\left(\frac{1-\beta^t}{1-\beta}\right)$ + ε_{it} + $\beta\varepsilon_{it-1}$ + $\beta^2\varepsilon_{it-1}$ + … +$\beta^{t-1}\varepsilon_{i1}$

= $\alpha_i\left(\frac{1-\beta^t}{1-\beta}\right)$ + $\sum_{s=1}^{t}\beta^{t-s}\,\varepsilon_{is}$

(iv) Compute E[$y_{it-1}(\alpha_i + \varepsilon_{it})$]. Is the OLS estimator of β applied to equation (1) consistent?

E[$y_{it-1}(\alpha_i + \varepsilon_{it})$] = E[$y_{it-1}$]E[$\alpha_i$] + E[$y_{it-1}$]E[$\varepsilon_{it}$] = E[$y_{it-1}\alpha_i$]

Which implies:

$y_{it-1} = y_{it-2}\beta + \alpha_i + \varepsilon_{it-1}$

No, unless we can eliminate α_i through transformation.

(v) Are the assumptions of the Random Effects model satisfied?

No, because if RE include lagged y_{it}, it will violate one of the RE's assumptions on exogeneity in which E[$\varepsilon_{it}|x_i$] = 0. If this assumption is violated, then the RE estimator cannot be consistent.

(vi) Apply the within-transformation to (1). Is the OLS estimator of β applied to the within transformed equation consistent?

No, the OLS estimator of β applied to the within transformed equation will not be consistent since:

$y_{it-1} = y_{it-2}\beta + \alpha_i + \varepsilon_{it-1}$

Then:

$\tilde{x}_{it} = y_{it-1} - \overline{y_\iota}$

$\tilde{\varepsilon}_{it} = \varepsilon_{it} - \overline{\varepsilon_\iota}$

E[$(y_{it-1} - \overline{y_\iota})(\varepsilon_{it} - \overline{\varepsilon_\iota})$] $\neq$ 0 (Inconsistent since strict exogeneity is not met).

(vii) Apply the first difference transformation to (1). Is the OLS estimator of β applied to the first differenced equation consistent?

No, it will not be consistent since:

$\tilde{x}_{it} = y_{it-1} - y_{it-2}$

$\tilde{\varepsilon}_{it} = \varepsilon_{it} - \varepsilon_{it-1}$

E[$(y_{it-1} - y_{it-2})(\varepsilon_{it} - \varepsilon_{it-1})$] $\neq$ 0 (Inconsistent)

(viii) Show that E[$y_{it-2}(\varepsilon_{it} - \varepsilon_{it-1})$] = 0. Propose a consistent estimator of β.

Firstly, to show that E[$y_{it-2}(\varepsilon_{it} - \varepsilon_{it-1})$] = 0:

E[$y_{it-2}(\varepsilon_{it} - \varepsilon_{it-1})$] = E[$y_{it-2}$]E[$\varepsilon_{it}$] - E[$y_{it-2}$]E[$\varepsilon_{it-1}$]

= E[y_{it-2}] $\times$ 0 - E[y_{it-2}] $\times$ 0 = 0

Secondly, to get a consistent estimator of β, we should incorporate the FD with instrumental variable. Initially let:

$\tilde{x}_{it} = y_{it-1} - y_{it-2}$

$\tilde{\varepsilon}_{it} = \varepsilon_{it} - \varepsilon_{it-1}$

Recall that moment condition for weak exogeneity:

$E[\varepsilon_{it}|\alpha_i, y_{i1},\ldots, y_{it-1}] = 0$

$E[\varepsilon_{it-1}|\alpha_i, y_{i1},\ldots, y_{it-2}] = 0$

Then incorporate the instrumental variable with FD:

$E[\tilde{z}_{it}\tilde{\varepsilon}_{it}] = 0$

Take $\tilde{z}_{it} = y_{it-2}$

$E[y_{it-2}(y_{it-1} - y_{it-2})] \neq 0$

Plim $\widehat{\beta_{IV}} = \beta + \frac{E[\tilde{z}_{it}\tilde{\varepsilon}_{it}]}{E[\tilde{z}_{it}\tilde{x}_{it}]} = \beta$

This is Arellano-Bond estimator in which:

Step 1: FD transformation

Step 2: IV with instrument y_{it-2}, y_{it-3}…so on.

Q4. Data Analysis

(i)

Can $\boldsymbol{\beta}$ be directly interpreted as a labor supply elasticity? Explain.

Since this is a log-log model, the coefficient β is the elasticity for the hourly wage and annual hours worked. We also call it as a constant elasticity model.

(ii)

The estimated coefficient on log wage $\widehat{\boldsymbol{\beta}}$, the default standard error and the standard error clustered on individual of the following estimators: population average, between, within, first-differences, and random effects.

(1) Population average

```
. //Population Average
. xtset id yr
       panel variable:  id (strongly balanced)
        time variable:  yr, 1979 to 1988
                delta:  1 unit

. xtreg lnhrs lnwg age agesq chld bdhlth, pa

Iteration 1: tolerance = .03117148
Iteration 2: tolerance = .00049767
Iteration 3: tolerance = 9.910e-06
Iteration 4: tolerance = 1.982e-07

GEE population-averaged model                  Number of obs     =      5,320
Group variable:                         id      Number of groups  =        532
Link:                             identity      Obs per group:
Family:                           Gaussian                    min =         10
Correlation:                  exchangeable                    avg =       10.0
                                                              max =         10
                                               Wald chi2(5)      =     100.86
Scale parameter:                  .0799038      Prob > chi2       =     0.0000
```

lnhrs	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnwg	.1165931	.0137213	8.50	0.000	.0896999	.1434863
age	.0077788	.0053887	1.44	0.149	-.0027828	.0183404
agesq	-.000102	.0000677	-1.51	0.132	-.0002348	.0000307
chld	.0046707	.0049376	0.95	0.344	-.0050068	.0143482
bdhlth	-.0692415	.0172447	-4.02	0.000	-.1030405	-.0354425
_cons	7.209251	.1042189	69.17	0.000	7.004986	7.413517

```
. estimates store PA

. xtreg lnhrs lnwg age agesq chld bdhlth, pa vce(bootstrap, reps(200))
(running xtgee on estimation sample)

Bootstrap replications (200)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200

GEE population-averaged model                   Number of obs     =      5,320
Group variable:                         id      Number of groups  =        532
Link:                             identity      Obs per group:
Family:                           Gaussian                    min =         10
Correlation:                  exchangeable                    avg =       10.0
                                                              max =         10
                                                Wald chi2(5)      =      18.82
Scale parameter:                  .0799038      Prob > chi2       =     0.0021

                                   (Replications based on 532 clusters in id)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
       lnhrs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |   .1165931   .0487795     2.39   0.017     .0209871    .2121992
         age |   .0077788   .0076535     1.02   0.309    -.0072219    .0227795
       agesq |   -.000102   .0001001    -1.02   0.308    -.0002982    .0000941
        chld |   .0046707   .0066854     0.70   0.485    -.0084324    .0177737
      bdhlth |  -.0692415   .0271277    -2.55   0.011    -.1224108   -.0160722
       _cons |   7.209251   .1882392    38.30   0.000     6.840309    7.578194
------------------------------------------------------------------------------

. estimates store PAC
```

(2) Between estimator

```
. //Between
. xtreg lnhrs lnwg age agesq chld bdhlth, be

Between regression (regression on group means)  Number of obs     =      5,320
Group variable: id                              Number of groups  =        532

R-sq:                                           Obs per group:
     within  = 0.0081                                         min =         10
     between = 0.0454                                         avg =       10.0
     overall = 0.0224                                         max =         10

                                                F(5,526)          =       5.01
sd(u_i + avg(e_i.))=  .1757244                  Prob > F          =     0.0002

------------------------------------------------------------------------------
       lnhrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |   .0658455     .02027     3.25   0.001     .0260253    .1056656
         age |  -.0075717   .0107101    -0.71   0.480    -.0286117    .0134682
       agesq |   .0000807   .0001316     0.61   0.540    -.0001778    .0003393
        chld |   .0081668   .0083201     0.98   0.327    -.0081778    .0245115
      bdhlth |  -.1390647   .0467947    -2.97   0.003    -.2309921   -.0471373
       _cons |   7.648018   .2061973    37.09   0.000     7.242947     8.05309
------------------------------------------------------------------------------
```

```
. estimates store BE

. xtreg lnhrs lnwg age agesq chld bdhlth, be vce(bootstrap, reps(200))
(running xtreg on estimation sample)

Bootstrap replications (200)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200

Between regression (regression on group means)  Number of obs     =      5,320
Group variable: id                              Number of groups  =        532

R-sq:                                           Obs per group:
     within  = 0.0081                                         min =         10
     between = 0.0454                                         avg =       10.0
     overall = 0.0224                                         max =         10

                                                Wald chi(5)       =      15.91
sd(u_i + avg(e_i.))=  .1757244                  Prob > chi2       =     0.0071

                                    (Replications based on 532 clusters in id)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
       lnhrs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |   .0658455   .0245266     2.68   0.007     .0177742    .1139168
         age |  -.0075717   .0095246    -0.79   0.427    -.0262396    .0110961
       agesq |   .0000807    .000119     0.68   0.497    -.0001525    .0003139
        chld |   .0081668   .0080891     1.01   0.313    -.0076876    .0240212
      bdhlth |  -.1390647   .0544211    -2.56   0.011    -.2457281   -.0324013
       _cons |   7.648018   .1909766    40.05   0.000     7.273711    8.022326
------------------------------------------------------------------------------

. estimates store BEC
```

(3) Within estimator

```
. //Within
. xtreg lnhrs lnwg age agesq chld bdhlth, fe

Fixed-effects (within) regression               Number of obs     =      5,320
Group variable: id                              Number of groups  =        532

R-sq:                                           Obs per group:
     within  = 0.0200                                         min =         10
     between = 0.0241                                         avg =       10.0
     overall = 0.0180                                         max =         10

                                                F(5,4783)         =      19.56
corr(u_i, Xb)  = -0.2102                        Prob > F          =     0.0000

------------------------------------------------------------------------------
       lnhrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |   .1647719   .0189379     8.70   0.000     .1276448    .2018989
         age |   .0142179   .0063804     2.23   0.026     .0017095    .0267263
       agesq |  -.0001676   .0000814    -2.06   0.039    -.0003272   -8.10e-06
        chld |  -.0011805   .0062019    -0.19   0.849    -.0133391    .0109781
      bdhlth |  -.0628007   .0185922    -3.38   0.001      -.09925   -.0263514
       _cons |   6.945659   .1258117    55.21   0.000      6.69901    7.192308
-------------+----------------------------------------------------------------
     sigma_u |  .18174487
     sigma_e |   .2324293
         rho |  .37943093   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(531, 4783) = 5.74                   Prob > F = 0.0000

. estimates store FE

. xtreg lnhrs lnwg age agesq chld bdhlth, fe vce(cluster id)

Fixed-effects (within) regression               Number of obs     =      5,320
Group variable: id                              Number of groups  =        532

R-sq:                                           Obs per group:
     within  = 0.0200                                         min =         10
     between = 0.0241                                         avg =       10.0
     overall = 0.0180                                         max =         10

                                                F(5,531)          =       2.71
corr(u_i, Xb)  = -0.2102                        Prob > F          =     0.0197

                                 (Std. Err. adjusted for 532 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lnhrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |   .1647719   .0863241     1.91   0.057    -.0048068    .3343505
         age |   .0142179   .0095997     1.48   0.139    -.0046402    .0330759
       agesq |  -.0001676   .0001258    -1.33   0.183    -.0004147    .0000794
        chld |  -.0011805     .00959    -0.12   0.902    -.0200195    .0176586
      bdhlth |  -.0628007   .0320411    -1.96   0.051    -.1257435    .0001421
       _cons |   6.945659   .2770255    25.07   0.000     6.401459    7.489859
-------------+----------------------------------------------------------------
     sigma_u |  .18174487
     sigma_e |   .2324293
         rho |  .37943093   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. estimates store FEC
```

(4) First-difference estimator

```
. //FD
. reg D.lnhrs D.(lnwg age agesq chld bdhlth), nocons

      Source |       SS           df       MS      Number of obs   =     4,788
-------------+----------------------------------   F(5, 4783)      =      6.68
       Model |  2.91492302         5  .582984605   Prob > F        =    0.0000
    Residual |  417.314267     4,783  .087249481   R-squared       =    0.0069
-------------+----------------------------------   Adj R-squared   =    0.0059
       Total |   420.22919     4,788  .087767166   Root MSE        =    .29538

------------------------------------------------------------------------------
     D.lnhrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |
         D1. |    .110343   .0213422     5.17   0.000     .0685026    .1521835
             |
         age |
         D1. |  -.0050943   .0195894    -0.26   0.795    -.0434985      .03331
             |
       agesq |
         D1. |   6.16e-06   .0002473     0.02   0.980    -.0004787     .000491
             |
        chld |
         D1. |  -.0098474   .0107984    -0.91   0.362    -.0310172    .0113225
             |
      bdhlth |
         D1. |  -.0421651   .0190354    -2.22   0.027    -.0794833   -.0048469
------------------------------------------------------------------------------

. estimates store FD

. reg D.lnhrs D.(lnwg age agesq chld bdhlth), nocons vce(cluster id)

Linear regression                               Number of obs     =      4,788
                                                F(5, 531)         =       0.86
                                                Prob > F          =     0.5046
                                                R-squared         =     0.0069
                                                Root MSE          =     .29538

                                  (Std. Err. adjusted for 532 clusters in id)
------------------------------------------------------------------------------
             |               Robust
     D.lnhrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |
         D1. |    .110343   .0838023     1.32   0.189    -.0542817    .2749678
             |
         age |
         D1. |  -.0050943   .0172881    -0.29   0.768    -.0390557    .0288672
             |
       agesq |
         D1. |   6.16e-06   .0001873     0.03   0.974    -.0003617     .000374
             |
        chld |
         D1. |  -.0098474   .0103429    -0.95   0.341    -.0301653    .0104706
             |
      bdhlth |
         D1. |  -.0421651   .0413417    -1.02   0.308    -.1233785    .0390483
------------------------------------------------------------------------------

. estimates store FDC
```

(5) Random effect estimator

```
. //RE
. xtreg lnhrs lnwg age agesq chld bdhlth, re

Random-effects GLS regression                   Number of obs     =      5,320
Group variable: id                              Number of groups  =        532

R-sq:                                           Obs per group:
     within  = 0.0188                                         min =         10
     between = 0.0327                                         avg =       10.0
     overall = 0.0222                                         max =         10

                                                Wald chi2(5)      =     100.78
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       lnhrs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |   .1163417   .0137023     8.49   0.000     .0894857    .1431977
         age |   .0077336   .0053867     1.44   0.151    -.0028241    .0182913
       agesq |  -.0001015   .0000677    -1.50   0.134    -.0002342    .0000311
        chld |   .0046998   .0049336     0.95   0.341    -.0049698    .0143694
      bdhlth |   -.069324   .0172511    -4.02   0.000    -.1031355   -.0355126
       _cons |   7.210815   .1041702    69.22   0.000     7.006645    7.414985
-------------+----------------------------------------------------------------
     sigma_u |  .15961431
     sigma_e |   .2324293
         rho |  .32046156   (fraction of variance due to u_i)
------------------------------------------------------------------------------


. estimates store RE

. xtreg lnhrs lnwg age agesq chld bdhlth, re vce(cluster id)

Random-effects GLS regression                   Number of obs     =      5,320
Group variable: id                              Number of groups  =        532

R-sq:                                           Obs per group:
     within  = 0.0188                                         min =         10
     between = 0.0327                                         avg =       10.0
     overall = 0.0222                                         max =         10

                                                Wald chi2(5)      =      16.22
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0062

                                  (Std. Err. adjusted for 532 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lnhrs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnwg |   .1163417   .0519834     2.24   0.025      .014456    .2182273
         age |   .0077336   .0079889     0.97   0.333    -.0079244    .0233916
       agesq |  -.0001015   .0001039    -0.98   0.328    -.0003052    .0001021
        chld |   .0046998   .0066227     0.71   0.478    -.0082805    .0176801
      bdhlth |   -.069324   .0285972    -2.42   0.015    -.1253736   -.0132744
       _cons |   7.210815   .1877884    38.40   0.000     6.842757    7.578874
-------------+----------------------------------------------------------------
     sigma_u |  .15961431
     sigma_e |   .2324293
         rho |  .32046156   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. estimates store REC
```

Summary table:

	PA	Between	Within	FD	RE
Coefficient on lnwg	0.1166	0.0659	0.1648	0.1103	0.1164
SE	0.0137	0.0203	0.0189	0.0213	0.0137
Clustered SE	0.0488	0.0245	0.0863	0.0838	0.0520

Comment:

As we can see from the summary table above, the within estimator generates the largest coefficient on while the between estimator results in the smallest coefficient on lnwg. More explanations will be discussed in part (iii). The clustered SE are consistently much larger than SE throughout all estimators and the reason behind this phenomenon will be explained in part (iv).

(iii)

Are the estimates of $\boldsymbol{\beta}$ similar? What could account for any differences you observe?

The coefficient estimates of β are different across different estimators. Several estimators used in this case involves eliminating α_i but each estimator has its own distinct feature. When RE estimator is applied if the RE model is appropriate, it will be inconsistent if the FE model is appropriate. Moreover, the between estimator is inconsistent in the FE model which allows for correlation between α_i and x_{it} but is consistent in the RE model which assumes that α_i is purely random and uncorrelated with the regressors. Another popular way to eliminate α_i is to use the FD estimator which has totally different sample scheme and has the vantage in relying on weaker exogeneity assumption that allows future values of the regressors to be correlated with the error. All of the disparities among the methodology of estimators result in the differences.

(iv)

Is there a systematic difference between the default and clustered standard errors for $\boldsymbol{\beta}$? What could account for any difference?

The clustered standard errors are much larger than the default standard errors for β. The clustered standard errors tend to reduce the precision for $\widehat{\beta}$ and caused the downward bias of variance of $\widehat{\beta}$ from the true variance. This phenomenon is due to the fact that the clustered standard errors allow for model errors to be correlated within cluster. In other words, model errors in different time periods for a given individual (e.g., person or firm or region) may be correlated, while model errors for different individuals are assumed to be uncorrelated. With the presence of clustered standard errors, OLS estimates remain unbiased but standard errors can be troublesome, leading to incorrect inference with finite samples.

(v)

Perform a Hausman test of the difference between the fixed effects and random effects estimators. What do you conclude? Which model is favored?

```
. //Hausman test
. hausman FE RE, sigmamore

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      FE           RE          Difference          S.E.
-------------+----------------------------------------------------------------
        lnwg |    .1647719     .1163417        .0484302        .0131157
         age |    .0142179     .0077336        .0064843        .0034381
       agesq |   -.0001676    -.0001015       -.0000661        .0000454
        chld |   -.0011805     .0046998       -.0058803        .0037742
      bdhlth |   -.0628007     -.069324        .0065233        .0070115
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(5) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       21.67
                Prob>chi2 =      0.0006
```

Conclusion from the Hausman test:

The overall statistics, $\chi^2(5)$ has p-value = 0.0006 which leads to strong rejection of the null hypothesis that RE provides consistent estimates. Thus, we favor the FE model.

(vi)

Suppose that $lnwg_{it}$ is endogenous due to correlation with both α_i and ε_{it}. Propose a consistent estimator of β. State any assumptions that you make.

I use Arellano-Bond estimator with additional regressor to deal with this problem. For consistent estimation, ε_{it} will not be serially correlated and can be tested easily on any econometric software. The Arellano-Bond estimator assumes E($y_{it}\Delta\varepsilon_{it}$) = 0 for s $\leq$ t $-$ 2. The $lnwg_{it}$ will become weakly exogenous in which the weak exogeneity assumption is shown as: E[$Z_{is}\varepsilon_{it}$] $= 0$ for s $\leq$ t that leads to E[$Z_{is}(\varepsilon_{it} - \varepsilon_{i,t-1})$] $= 0$. The instruments are the first differenced regressor x_{it} and the second lag of $lnwg_{it}$. For s $\leq$ t -1, this method will reduce the available instrument set by one period. More clearly, $\Delta\varepsilon_{it}$ = $\varepsilon_{it} - \varepsilon_{it-1}$ is correlated with $\Delta y_{i,t-1}$= $y_{i,t-1} - y_{i,t-2}$ because $y_{i,t-1}$ depends on ε_{it-1} . Further, $\Delta\varepsilon_{it}$ is uncorrelated with $\Delta y_{i,t-k}$ for k $\geq$ 2. Altogether, this transformation will result in consistent estimator of β.

(vii)

Implement your estimator from (vi) and interpret the results.

Arellano-Bond test for zero autocorrelation in FD errors:

Order	z	Prob > z
1	-3.8765	0.0001
2	.97517	0.3295
3	.5608	0.5749

H0: no autocorrelation

Arellano-Bond dynamic panel data model:

```
Arellano-Bond dynamic panel-data estimation     Number of obs      =      3,724
Group variable: id                              Number of groups   =        532
Time variable: yr
                                                Obs per group:
                                                              min =          7
                                                              avg =          7
                                                              max =          7

Number of instruments =     47                  Wald chi2(9)       =      41.76
                                                Prob > chi2        =     0.0000
Two-step results
                                     (Std. Err. adjusted for clustering on id)
```

lnhrs	Coef.	WC-Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnhrs						
L1.	.202109	.0882424	2.29	0.022	.0291571	.375061
L2.	.0948946	.0468263	2.03	0.043	.0031168	.1866725
lnwg						
--.	.953449	.3083689	3.09	0.002	.349057	1.557841
L1.	-.4737369	.3286299	-1.44	0.149	-1.11784	.1703658
L2.	.0006799	.0812386	0.01	0.993	-.1585449	.1599047
age	.0013365	.0190764	0.07	0.944	-.0360525	.0387255
agesq	1.50e-06	.0002274	0.01	0.995	-.0004441	.0004471
bdhlth	-.0478328	.0336125	-1.42	0.155	-.1137121	.0180465
chld	-.0176063	.0155253	-1.13	0.257	-.0480353	.0128227
_cons	4.098712	1.463389	2.80	0.005	1.230523	6.966901

Comment:

Firstly, the Arellano-Bond test for zero autocorrelation in FD errors shows that the serial correlation can be eliminated at order two; therefore, the corresponding model should have at least two lags on lnhrs and lnwg. As we can see from the regression table above, lnwg, L1.lnwg, and L2.lnwg are regressors and only four additional lags are sets as instruments. The max lags that are allowed to be used for the dependent variable are set at two. The impact of lnwg on lnhrs will be greatly reduced at lag two and the impact of lnhrs lag two on itself will also be greatly reduced and is statistically significant at five percent level.

Appendix: STATA codes

```
clear
capture log close
set logtype text
log using ECONA2CSLO.txt, replace

cd "D:\UQ\2020 S1\ECON7320 Advanced MicroMetrics\Assignment 2"
use MOM.dta

//Population Average
xtset id yr
xtreg lnhrs lnwg age agesq chld bdhlth, pa
estimates store PA
xtreg lnhrs lnwg age agesq chld bdhlth, pa vce(bootstrap, reps(200))
estimates store PAC

//Between
xtreg lnhrs lnwg age agesq chld bdhlth, be
estimates store BE
xtreg lnhrs lnwg age agesq chld bdhlth, be vce(bootstrap, reps(200))
estimates store BEC

//Within
xtreg lnhrs lnwg age agesq chld bdhlth, fe
estimates store FE
xtreg lnhrs lnwg age agesq chld bdhlth, fe vce(cluster id)
estimates store FEC

//FD
reg D.lnhrs D.(lnwg age agesq chld bdhlth), nocons
estimates store FD
reg D.lnhrs D.(lnwg age agesq chld bdhlth), nocons vce(cluster id)
```

```
estimates store FDC

//RE
xtreg lnhrs lnwg age agesq chld bdhlth, re
estimates store RE
xtreg lnhrs lnwg age agesq chld bdhlth, re vce(cluster id)
estimates store REC

//Hausman test
hausman FE RE, sigmamore

//Arellano-Bond estimator (ABE)
xtabond lnhrs age agesq bdhlth chld,lags(2) maxldep(3) pre(lnwg, lag(2,4)) twostep vce(robust)
artests(3)

//ABE seriall correlationt est//
estat abond

log cl
```