BACHELOR OF ECONOMICS
THAMMASAT UNIVERSITY

Answer Sheet Cover Page
Midterm Examination Semester 2/2020

(Readable handwriting and printed version are acceptable)

Student Name: Harit Hatawong

Student ID: 6104640112

Course ID: EE426 Course Title: Econometrics 2

Lecturer: Assoc. Prof. Tatre Jantarakolica

Exam date: 21/03/2021 Time: 9.00 – 15.00

Total pages: 22

***Students are responsible for checking the number of pages before submitting the answer sheets. Incomplete submissions caused by carelessness will not be accepted as an excuse for resubmission**

Student Signature: Harit Hatawong

Date: 21/03/2021

1. From the data set "Midterm_q1_no.dta": Estimate the following models

   a. Estimate model (1) and (2) using Ordinary Least Squares (OLS) and state consequences of using OLS in this case (5 Points)

```
.reg y1 y2 x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(2, 47)        =    105.95
       Model |  51488.9634         2  25744.4817   Prob > F        =    0.0000
    Residual |  11420.2566        47  242.984184   R-squared       =    0.8185
-------------+----------------------------------   Adj R-squared   =    0.8107
       Total |    62909.22        49  1283.86163   Root MSE        =    15.588

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          y2 |   .0009133   .0001911     4.78   0.000     .0005289    .0012977
          x3 |   .0075035   .0010848     6.92   0.000     .0053212    .0096858
       _cons |   43.08226   16.70979     2.58   0.013     9.466507    76.69801
------------------------------------------------------------------------------

. reg y2 y1 x1 x2

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     30.18
       Model |  8.1086e+09         3  2.7029e+09   Prob > F        =    0.0000
    Residual |  4.1191e+09        46  89546305.3   R-squared       =    0.6631
-------------+----------------------------------   Adj R-squared   =    0.6412
       Total |  1.2228e+10        49   249545016   Root MSE        =    9462.9

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          y1 |   302.2265   47.05593     6.42   0.000     207.5078    396.9452
          x1 |   253.5806   133.7739     1.90   0.064    -15.69221    522.8534
          x2 |  -.0004903   .0003309    -1.48   0.145    -.0011563    .0001757
       _cons |  -36983.86   9157.432    -4.04   0.000    -55416.81    -18550.9
------------------------------------------------------------------------------
```

The consequences of using OLS in this case is the model will be biased since the y2 is correlate with the error term in model 1 and the y1 is also correlate with the error term in model2 so, the endogeneity bias occurred.

   b. Estimate model (1) and (2) using Two Stage Least Squares (2SLS) and state reduced form and structural form of these simultaneous equation models. Specify endogenous variables and exogenous variables. Then, estimate reduced form models using OLS and structural form models using IV technique from the predicted endogenous variables from reduced form estimated results. (7 points)

$$y_{1t} = \beta_{10} + \gamma_{12} y_{2t} + \beta_{13} x_{3t} + u_{1t} \quad (1)$$

$$y_{2t} = \beta_{20} + \gamma_{21} y_{1t} + \beta_{21} x_{1t} + \beta_{22} x_{2t} + u_{2t} \quad (2)$$

substitute (1) in (2)

$$Y_{2t} = \beta_{20} + \gamma_{21}(\beta_{10} + \gamma_{12} y_{2t} + \beta_{13} x_{3t} + u_{1t}) + \beta_{21} X_{1t} + \beta_{22} X_{2t} + U_{2t}$$

$$\therefore Y_{2t} = \Pi_{20} + \Pi_{21} X_{1t} + \Pi_{22} X_{2t} + \Pi_{23} X_{3t} + W_{2t}$$

substitute (2) in (1)

$$Y_{1t} = \beta_{10} + \gamma_{12}(\beta_{20} + \gamma_{21} y_{1t} + \beta_{21} x_{1t} + \beta_{22} x_{2t} + u_{2t}) + \beta_{13} X_{3t} + U_{1t}$$

$$\therefore Y_{1t} = \Pi_{10} + \Pi_{11} X_{1t} + \Pi_{12} X_{2t} + \Pi_{13} X_{3t} + W_{1t}$$

reduced form

$$Y_{1t} = \Pi_{10} + \Pi_{11} X_{1t} + \Pi_{12} X_{2t} + \Pi_{13} X_{3t} + W_{1t}$$

$$Y_{2t} = \Pi_{20} + \Pi_{21} X_{1t} + \Pi_{22} X_{2t} + \Pi_{23} X_{3t} + W_{2t}$$

structural form

$$y_{1t} = \beta_{10} + \gamma_{12} \hat{y}_{2t} + \beta_{13} x_{3t} + u_{1t}$$

$$y_{2t} = \beta_{20} + \gamma_{21} \hat{y}_{1t} + \beta_{21} x_{1t} + \beta_{22} x_{2t} + u_{2t}$$

Endogenous variable: y1 y2
Exogeneous variable: x1 x2 x3

Estimated reduced form

```
. reg y1 x1 x2 x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     45.00
       Model |  46922.5555         3  15640.8518   Prob > F        =    0.0000
    Residual |  15986.6645        46  347.536184   R-squared       =    0.7459
-------------+----------------------------------   Adj R-squared   =    0.7293
       Total |    62909.22        49  1283.86163   Root MSE        =    18.642

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .4333597   .2592293     1.67   0.101    -.0884419    .9551612
          x2 |  -3.99e-07   6.49e-07    -0.61   0.542    -1.71e-06    9.08e-07
          x3 |   .0098608   .0011755     8.39   0.000     .0074945     .012227
       _cons |   14.16024    19.2523     0.74   0.466    -24.59264    52.91311
------------------------------------------------------------------------------

. predict y1hat
(option xb assumed; fitted values)

. reg y2 x1 x2 x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     17.04
       Model |  6.4355e+09         3  2.1452e+09   Prob > F        =    0.0000
    Residual |  5.7922e+09        46   125918002   R-squared       =    0.5263
-------------+----------------------------------   Adj R-squared   =    0.4954
       Total |  1.2228e+10        49   249545016   Root MSE        =     11221

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   402.7605   156.0372     2.58   0.013     88.67401     716.847
          x2 |  -.0006235   .0003909    -1.60   0.118    -.0014104    .0001633
          x3 |   2.834631   .7075866     4.01   0.000     1.410333    4.258929
       _cons |  -30998.27   11588.48    -2.67   0.010    -54324.67    -7671.86
------------------------------------------------------------------------------

. predict y2hat
(option xb assumed; fitted values)
```

Estimated structural form

```
. reg y1 y2hat x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(2, 47)        =     68.54
       Model |  46846.5903         2  23423.2952   Prob > F        =    0.0000
    Residual |  16062.6297        47  341.758078   R-squared       =    0.7447
-------------+----------------------------------   Adj R-squared   =    0.7338
       Total |    62909.22        49  1283.86163   Root MSE        =    18.487

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y2hat |   .0010261   .0006294     1.63   0.110    -.0002402    .0022923
          x3 |   .0070711   .0025923     2.73   0.009     .0018561    .0122861
       _cons |   46.07603   25.21034     1.83   0.074    -4.640639    96.79269
------------------------------------------------------------------------------
```

```
. reg y2 y1hat x1 x2

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     17.04
       Model |  6.4355e+09         3  2.1452e+09   Prob > F        =    0.0000
    Residual |  5.7922e+09        46   125918006   R-squared       =    0.5263
-------------+----------------------------------   Adj R-squared   =    0.4954
       Total |  1.2228e+10        49   249545016   Root MSE        =     11221

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y1hat |   287.4652   71.75768     4.01   0.000     143.0245    431.9059
          x1 |   278.1847   175.5543     1.58   0.120    -75.18772    631.5571
          x2 |  -.0005088   .0003964    -1.28   0.206    -.0013068    .0002891
       _cons |  -35068.84   12336.08    -2.84   0.007    -59900.09   -10237.59
------------------------------------------------------------------------------
```

c. Estimate model (1) and (2) using Three Stage Least Squares (3SLS) and give the explanation of the differences among the three estimation methods (OLS, 2SLS, and 3SLS) (conceptually). Point out the advantage and disadvantage in term of properties of estimated results from each method (single equation estimation vs system equations estimation methods). (8 points)

<u>3SLS Estimated model</u>

```
. reg3 (y1 y2 x3) (y2 y1 x1 x2), 3sls inst(x1 x2 x3)

Three-stage least-squares regression
----------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
y1                    50       2    15.16901    0.8171     203.59   0.0000
y2                    50       3    9100.904    0.6613      78.10   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0010261   .0005165     1.99   0.047     .0000138    .0020383
          x3 |   .0070711   .0021271     3.32   0.001     .0029022    .0112401
       _cons |   46.07603   20.68601     2.23   0.026     5.532198    86.61985
-------------+----------------------------------------------------------------
y2           |
          y1 |   283.4667   57.68578     4.91   0.000     170.4046    396.5287
          x1 |     286.58   141.3979     2.03   0.043     9.445245    563.7147
          x2 |  -.0004306   .0002907    -1.48   0.139    -.0010003    .0001391
       _cons |  -35048.35   9988.778    -3.51   0.000       -54626   -15470.71
------------------------------------------------------------------------------
Endogenous variables:  y1 y2
Exogenous variables:   x1 x2 x3
------------------------------------------------------------------------------
```

For the differences between 3 models, OLS, 2SLS, 3SLS in the case of endogeneity bias in simultaneous equation models.
OLS will give the biased, inconsistent and inefficient results
2SLS using the instrumental variable concept to fix the problem which the results are biased, consistent and asymptotically efficient
3SLS is the 2SLS + SUR which it should be use in the case that we have correlation between 2 model's error terms. For the results 3SLS give us biased, consistent and
MORE asymptotically efficient however, if we have specification error the impact will be severed in 3SLS.

2. From the data set "Final_q2_no.dta":
The study of cost function assumes the model follows CES cost function:

a. Estimate the model (4) using NLS estimation method using initial values of lnλ=1, θ=0.5, β=0.5, α= –0.5. Determine the estimated value of efficiency parameter (λ), distribution parameter (θ), parameter (β), and substitution parameter (α), and elasticity of substitution (σ). Perform F-test to test whether θ=0, α=0, and β=0. (6 points)

Estimated model with NLS

```
. nl ( lnC = {lnlamda} - ({beta}/{alpha})*ln({zeta}*(R^(-{alpha})) + (1-{zeta})*(W^(-{al
> pa})))), init(lnlamda 1 zeta 0.5 beta 0.5 alpha -0.5)
(obs = 252)

Iteration 0:  residual SS =  2199.272
Iteration 1:  residual SS =  292.7041
...
Iteration 51:  residual SS =  279.1644


      Source |       SS            df       MS
-------------+----------------------------------    Number of obs =        252
       Model |   47.23926           4  11.8098151    R-squared     =     0.1447
    Residual |  279.16439         247  1.13022019    Adj R-squared =     0.1309
-------------+----------------------------------    Root MSE      =   1.063118
       Total |  326.40365         251  1.30041293    Res. dev.     =   740.9427

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    /lnlamda |   1.733219   1.014233     1.71   0.089    -.2644299    3.730868
       /beta |   .7257297   .1451206     5.00   0.000      .439898    1.011561
      /alpha |  -2.719751   9.974943    -0.27   0.785    -22.36655    16.92704
       /zeta |   .9999984   .0000726 13769.91   0.000     .9998554    1.000141
       /alpa |  -6.036286   22.90198    -0.26   0.792    -51.14436    39.07179
------------------------------------------------------------------------------
  Parameter lnlamda taken as constant term in model & ANOVA table
```

Estimated value of parameter

$\ln\lambda = 1.733219$ then $\lambda = 5.66$	which λ = lamda
$\theta = 0.9999984$	which θ = zeta
$\beta = 0.7257297$	which β = beta
$\alpha = -6.036286$	which α = alpha

$\sigma = \frac{1}{1-\theta} = 625{,}000$

F-test H0: θ=0, α=0, and β=0

```
.  test (_b[/zeta]=0) (_b[/alpha]=0) (_b[/beta]=0)

 ( 1)  [zeta]_cons = 0
 ( 2)  [alpha]_cons = 0
 ( 3)  [beta]_cons = 0

       F(  3,   247) = 2.3e+09
            Prob > F =    0.0000
```

Reject H0, we don't have evidence that θ=0, α=0, and β=0

a. What will happy if we change initial values to lnλ=0.5, θ=0.1, β=0.1, α= –0.1? Will the estimated results be the same as (a)? What are the differences between the previous result in (a) and the new result? Give explanation why? (6 points)

The results would not be the same since we change the initial values in the nonlinear model it's possibly that we will have more than 1 minimum point which mean that the root of nls will have more than 1 answer. So, if we change the initial value it will go to the closest minimum point which may different from the previous one with different initial value.

Estimated model with new initital value

```
. nl ( lnC = {lnlamda} - ({beta}/{alpha})*ln(({zeta}*(R^(-{alpha}))) + ((1-
{zeta})*(W^(-
> {alpa}))))), init(lnlamda 0.5 zeta 0.1 beta 0.1 alpha -0.1)
(obs = 252)

Iteration 0:  residual SS =  9967.773
Iteration 1:  residual SS =  9967.681
...
Iteration 49:  cannot calculate derivatives
equation/system not identified
r(481);
```

the differences is we get no answer in this case, it's possibly because of we set the initial value too far from the closest minimum point and impossible to get there with current convergence value.

b. From (b), if we change convergence value from default of 0.00001 or (1e-5) to (i) 0.1 or (1e-1) and (ii) (1e-15) with maximum iteration of 40, what will happen to the estimated result? Interpret the estimated result and why do we get this kind of result? (Make comparison between previous result in (b) and the new result) (6 points)

According to my answer in (b) , the previous case with the convergence value = 1e-5 it impossible to get to the closest root so, if we increase convergence value to 1e-1 it's possibly that we will get the answer and for the 1e-15 with iter(40) I think it's also possible to get the answer as well but the convergence will not be achieved for sure.

NLS with 1e-1 convergence

```
. nl ( lnC = {lnlamda} - ({beta}/{alpha})*ln(({zeta}*(R^(-{alpha}))) + ((1-{zeta})*(W^(-{
> alpa}))))), init(lnlamda 0.5 zeta 0.1 beta 0.1 alpha -0.1) eps(1e-1)
(obs = 252)

Iteration 0:  residual SS =  9967.773


      Source |      SS            df       MS
-------------+----------------------------------    Number of obs =        252
       Model | -9641.3692          4  -2410.3423    R-squared     =   -29.5382
    Residual |  9967.7728        247  40.3553556    Adj R-squared =   -30.0327
-------------+----------------------------------    Root MSE      =   6.352587
       Total |  326.40365        251  1.30041293    Res. dev.     =   1641.921

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    /lnlamda |   .5000699   1.754947     0.28   0.776    -2.956499    3.956639
       /beta |   .0972794   558.1485     0.00   1.000     -1099.24    1099.435
      /alpha |  -.1006157   103.6958    -0.00   0.999    -204.3413    204.1401
       /zeta |   .1026577   548.8484     0.00   1.000    -1080.917    1081.123
       /alpa |  -3.66e-06   .0801858    -0.00   1.000    -.1579389    .1579316
------------------------------------------------------------------------------
  Parameter lnlamda taken as constant term in model & ANOVA table
```

NLS with 1e-15 convergence and max iterative = 40

```
. nl ( lnC = {lnlamda} - ({beta}/{alpha})*ln(({zeta}*(R^(-{alpha}))) + ((1-{zeta})*(W^(-{
> alpa}))))), init(lnlamda 0.5 zeta 0.1 beta 0.1 alpha -0.1) eps(1e-15) iter(40)
(obs = 252)

Iteration 0:  residual SS =  9967.773
Iteration 1:  residual SS =  9967.681
...
Iteration 39:  residual SS =  415.9504


      Source |       SS             df       MS
-------------+----------------------------------    Number of obs =        252
       Model | -89.546721            4 -22.3866803    R-squared     =    -0.2743
    Residual |  415.95037          247  1.68400959    Adj R-squared =    -0.2950
-------------+----------------------------------    Root MSE      =   1.297694
       Total |  326.40365          251  1.30041293    Res. dev.     =   841.4315

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    /lnlamda |   1.754187   .6058545     2.90   0.004     .5608868    2.947487
       /beta |   .7224001   .0892363     8.10   0.000     .5466389    .8981612
      /alpha |   13.11296   871.1872     0.02   0.988     -1702.79    1729.016
       /zeta |   .9998336   .0333679    29.96   0.000     .9341117    1.065556
       /alpa |   23.71051   1662.979     0.01   0.989    -3251.718    3299.139
------------------------------------------------------------------------------
  Parameter lnlamda taken as constant term in model & ANOVA table
convergence not achieved
```

from the Stata printout it can confirmed my assumption in both cases.

c. Why do we prefer to estimate nonlinear regression model in log-form instead of its original functional form? (2 points)

Because it will shorten the length of the data which making the process to be faster
As well as in the MLE models.

3. From the data set "Midterm_q3_no.dta": (using do-file from assignment 4)

a. Estimate the above models using MLE with (i) Newton-Ralphson algorithm; (ii) Berndt-Hall-Hall-Hausman algorithm; and (iii) Broyden-Fletcher-Goldfarb- Shanno algorithm, make comparison of the estimated results using different algorithm, and give explanation why are they different? (5 points)

Newton-Ralphson algorithm

```
. ml model lf ml_logit (y= x1 x2 )

. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood = -279.13079
rescale:       log likelihood = -275.12577
Iteration 0:   log likelihood = -275.12577
Iteration 1:   log likelihood = -227.96269
Iteration 2:   log likelihood = -227.67192
Iteration 3:   log likelihood = -227.67158
Iteration 4:   log likelihood = -227.67158

                                                Number of obs     =        400
                                                Wald chi2(2)      =      66.43
Log likelihood = -227.67158                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .2508771   .1098901     2.28   0.022     .0354965    .4662577
          x2 |  -.5626563   .0706508    -7.96   0.000    -.7011294   -.4241832
       _cons |   .6249761   .1975087     3.16   0.002     .2378661    1.012086
------------------------------------------------------------------------------
```

Berndt-Hall-Hall-Hausman algorithm

```
. ml model lf ml_logit (y= x1 x2 ), tech(bhhh)

. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood = -279.13079
rescale:       log likelihood = -275.12577
Iteration 0:   log likelihood = -275.12577
Iteration 1:   log likelihood = -227.90673
Iteration 2:   log likelihood = -227.67573
Iteration 3:   log likelihood = -227.67164
Iteration 4:   log likelihood = -227.67159
Iteration 5:   log likelihood = -227.67158

                                                Number of obs     =        400
                                                Wald chi2(2)      =      81.09
Log likelihood = -227.67158                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |                 OPG
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .2508881   .1109659     2.26   0.024      .033399    .4683772
          x2 |  -.5626591   .0646113    -8.71   0.000    -.6892949   -.4360233
       _cons |   .6249569   .2014597     3.10   0.002     .2301032    1.019811
------------------------------------------------------------------------------
```

Broyden-Fletcher-Goldfarb- Shanno algorithm

```
ml model lf ml_logit (y= x1 x2 ), tech(bfgs)

. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood = -279.13079
rescale:       log likelihood = -275.12577
Iteration 0:   log likelihood = -275.12577
Iteration 1:   log likelihood = -252.09628  (backed up)
Iteration 2:   log likelihood = -233.07538  (backed up)
Iteration 3:   log likelihood = -229.55938
Iteration 4:   log likelihood = -227.71884
Iteration 5:   log likelihood = -227.67316
Iteration 6:   log likelihood =  -227.6716
Iteration 7:   log likelihood = -227.67158

                                                Number of obs     =        400
                                                Wald chi2(2)      =      66.43
Log likelihood = -227.67158                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .2508827   .1098903     2.28   0.022     .0355017    .4662636
          x2 |  -.5626583    .070651    -7.96   0.000    -.7011317   -.4241849
       _cons |   .6249707   .1975088     3.16   0.002     .2378605    1.012081
------------------------------------------------------------------------------
```

in different algorithm we get the difference in estimated value because in each algorithm using different way to come up with the result some of them are using hessian matrix some of them are using gradient matrix so, it's normal to get different value from different method by the way it's not a huge different.

b. Perform hypothesis testing whether $B_1 = B_2 = 0$ using LR-test and Wald test. Make comparison between the two tests. Which test is preferable? Why? (5 points)

LR-test

```
.lrtest unres res

Likelihood-ratio test                                 LR chi2(2)  =     94.76
(Assumption: res nested in unres)                     Prob > chi2 =    0.0000
```

Wald test

```
.test x1 x2

 ( 1)  [eq1]x1 = 0
 ( 2)  [eq1]x2 = 0

           chi2(  2) =   66.43
         Prob > chi2 =    0.0000
```

We reject th H0 in both cases which mean that the model is jointly significant in overall test. However if the test results are different we will choose the LR-test result because it's more accurate for the Wald test the test depends on parameterization.

c. Why overall test in MLE is Chi-square test instead of F-test? Can we still employ F-test as overall test? Why or why not? (2 points)

For MLE I think we can use F-test instead of chi -square as well

d. Estimate the models with heteroskedasticity using MLE with Newton-Ralphson algorithm. Perform LR-test whether there exists significant heteroskedasticity. In this case, can we perform Wald-test or LM-test to test heteroskedasticity? Why? or why not? (5 points)

Estimate model with heteroskedasticity

```
. ml model lf ml_probit_het ( y = x1 x2) (x3, noconstant)

. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood =  -291.8231
rescale:       log likelihood = -275.03884
rescale eq:    log likelihood = -271.46187
Iteration 0:   log likelihood = -271.46187  (not concave)
Iteration 1:   log likelihood = -259.02902  (not concave)
…
Iteration 9:   log likelihood = -225.58236

                                                Number of obs     =        400
                                                Wald chi2(2)      =      64.18
Log likelihood = -225.58236                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
eq1          |
          x1 |   .1580213   .0626383     2.52   0.012     .0352526    .2807901
          x2 |  -.3231145   .0414919    -7.79   0.000    -.4044373   -.2417918
       _cons |   .3665324   .1132645     3.24   0.001     .1445381    .5885268
-------------+----------------------------------------------------------------
eq2          |
          x3 |   .3778106   .1707856     2.21   0.027      .043077    .7125441
------------------------------------------------------------------------------

. est store hetprobit
. ml model lf ml_probit ( y = x1 x2)
. ml maximize
. est store probit

. lrtest hetprobit probit

Likelihood-ratio test                                 LR chi2(1)  =      6.03
(Assumption: probit nested in hetprobit)              Prob > chi2 =    0.0141
```

In this case we reject the H0 which mean that the heteroskedasticity might be exist.

We can use wald test or LM as well but the lr test would give us the best result.

e. Why individual test in MLE is *z*-test instead of *t*-test? Why don't we have R- square in MLE? If we would like to make comparison between the two estimated results, which statistical index should be employed? (3 points)

For MLE we are using z-test since we assume asymptotically normal.

4.From the data set "Midterm_q4-1_no.dta":
The model of interest rate structure, continuous time model can be specified as:

a. Estimate Unrestricted model using method of moment, Merton model using GMM, and Vasicek using GMM. Make evaluation of the estimated result of Merton model. (5 points).

Unrestricted model(MOM)

```
. . gmm (dr-{alpha}-{beta}*r) ((dr-{alpha}-{beta}*r)*r) ((dr-{alpha}-{beta}*r)^2- {
> sigma2}*r^(2*{gamma})) (((dr-{alpha}-{beta}*r)^2-{sigma2}*r^(2*{gamma}))*r) winit
> ial(identity)
note: 1 missing value returned for equation 1 at initial values
note: 1 missing value returned for equation 2 at initial values
note: 1 missing value returned for equation 3 at initial values
note: 1 missing value returned for equation 4 at initial values

Step 1
numerical derivatives are approximate
flat or discontinuous region encountered
Iteration 0:   GMM criterion Q(b) =  .00001191
Iteration 1:   GMM criterion Q(b) =  8.722e-06  (backed up)
Iteration 2:   GMM criterion Q(b) =  6.129e-06  (not concave)
Iteration 3:   GMM criterion Q(b) =  5.573e-06  (backed up)
Iteration 4:   GMM criterion Q(b) =  5.489e-06  (backed up)

Step 2
Iteration 0:   GMM criterion Q(b) =   .0076731
Iteration 1:   GMM criterion Q(b) =   .0072924
Iteration 2:   GMM criterion Q(b) =  .00453733
Iteration 3:   GMM criterion Q(b) =  .00151095
Iteration 4:   GMM criterion Q(b) =   .0006721
Iteration 5:   GMM criterion Q(b) =  6.065e-07
Iteration 6:   GMM criterion Q(b) =  2.058e-13
Iteration 7:   GMM criterion Q(b) =  1.852e-24

note: model is exactly identified

GMM estimation

Number of parameters =   4
Number of moments    =   4
Initial weight matrix: Identity                   Number of obs   =      1,325
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /alpha |  -.0023917   .0011635    -2.06   0.040    -.0046722   -.0001112
       /beta |   .0004321   .0002872     1.50   0.132    -.0001309    .0009951
     /sigma2 |   .0005133   .0003281     1.56   0.118    -.0001299    .0011564
      /gamma |   .0942137   .1807945     0.52   0.602    -.2601371    .4485645
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
Instruments for equation 4: _cons
```

Merton model(GMM)

```
gmm (dr-{alpha}) ((dr-{alpha})*r) ((dr-{alpha})^2-{sigma2}) (((dr-{alpha})^2- {si
> gma2})*r) winitial(identity)
note: 1 missing value returned for equation 1 at initial values
note: 1 missing value returned for equation 2 at initial values
note: 1 missing value returned for equation 3 at initial values
note: 1 missing value returned for equation 4 at initial values

Step 1
Iteration 0:   GMM criterion Q(b) =  .00001191
Iteration 1:   GMM criterion Q(b) =  4.144e-08
Iteration 2:   GMM criterion Q(b) =  4.144e-08

Step 2
Iteration 0:   GMM criterion Q(b) =  .00793798
Iteration 1:   GMM criterion Q(b) =  .00554765
Iteration 2:   GMM criterion Q(b) =  .00554765

GMM estimation

Number of parameters =   2
Number of moments    =   4
Initial weight matrix: Identity                   Number of obs   =      1,325
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /alpha |  -.0008227   .0006877    -1.20   0.232    -.0021706    .0005252
     /sigma2 |   .0004354   .0002932     1.49   0.137    -.0001392      .00101
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
Instruments for equation 4: _cons

. estat overid

  Test of overidentifying restriction:

  Hansen's J chi2(2) = 7.35063 (p = 0.0253)
```

Vasicek model(GMM)

```
gmm (dr-{alpha}-{beta}*r) ((dr-{alpha}-{beta}*r)*r) ((dr-{alpha}-{beta}*r)^2- {s
> igma2}) (((dr-{alpha}-{beta}*r)^2-{sigma2})*r) winitial(identity)
note: 1 missing value returned for equation 1 at initial values
note: 1 missing value returned for equation 2 at initial values
note: 1 missing value returned for equation 3 at initial values
note: 1 missing value returned for equation 4 at initial values

Step 1
Iteration 0:   GMM criterion Q(b) =  .00001191
Iteration 1:   GMM criterion Q(b) =  3.145e-10
Iteration 2:   GMM criterion Q(b) =  3.079e-10

Step 2
Iteration 0:   GMM criterion Q(b) =   .0005879
Iteration 1:   GMM criterion Q(b) =  .00019279
Iteration 2:   GMM criterion Q(b) =  .00019279

GMM estimation

Number of parameters =   3
Number of moments    =   4
Initial weight matrix: Identity                   Number of obs   =      1,325
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /alpha |  -.0027101   .0009776    -2.77   0.006    -.0046262    -.000794
       /beta |   .0005363   .0001997     2.69   0.007     .0001449    .0009277
     /sigma2 |   .0005959   .0003005     1.98   0.047     6.90e-06     .001185
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
Instruments for equation 4: _cons

.
. estat overid

  Test of overidentifying restriction:

  Hansen's J chi2(1) = .255441 (p = 0.6133)
```

For the evaluation of merton model

1. Sign and meaning are normal
2. overall test from LR-test we can reject the H0 so, the model is jointly significant
3. for individual test we have insignificant in both parameters

b. Determine which model is the most appropriated model. Provide and give explanation of your selected criteria. (5 points)

Merton model wald test

```
test (_b[/beta]=0) (_b[/gamma]=0)

 ( 1)  [beta]_cons = 0
 ( 2)  [gamma]_cons = 0

           chi2(  2) =    7.89
         Prob > chi2 =    0.0194
```

Vasicek model wald test

```
. test (_b[/gamma]=0)

 ( 1)  [gamma]_cons = 0

           chi2(  1) =    0.27
         Prob > chi2 =    0.6023
```

According to the wald test vasicek would be the best model since we cannot reject the H0 which mean that gamma is equal to 0.

From the data set "`Midterm_q4-2_no.dta`":

c. What will happen to the estimated results if we estimate this model (10) using OLS? (2 points)

If we run the model with OLS the endogeneity bias will be occurred.

d. Estimate the model (10) using GMM and 2SLS by employing z_{1i}, z_{2i}, and z_{3i} as instrumental variables for x_{1i} and x_{2i}. Perform the test to check whether GMM is appropriated. (5 points).

Estimated model with GMM

```
ivregress gmm y x3 x4 (x1 x2 = z1 z2 z3)

Instrumental variables (GMM) regression          Number of obs   =        500
                                                 Wald chi2(4)    =     162.13
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.6096
GMM weight matrix: Robust                        Root MSE        =     19.919

------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   3.301395   .9218627     3.58   0.000     1.494578    5.108213
          x2 |   2.046575   .5101686     4.01   0.000     1.046663    3.046487
          x3 |   1.007346   .3058432     3.29   0.001     .4079044    1.606788
          x4 |   1.321124   .1736521     7.61   0.000     .9807718    1.661475
       _cons |   1.509203   7.597793     0.20   0.843    -13.3822     16.4006
------------------------------------------------------------------------------
Instrumented:  x1 x2
Instruments:   x3 x4 z1 z2 z3
```

Estimated model with 2SLS

```
. ivregress 2sls y x3 x4 (x1 x2 = z1 z2 z3)

Instrumental variables (2SLS) regression         Number of obs   =        500
                                                 Wald chi2(4)    =     153.63
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.6097
                                                 Root MSE        =     19.916

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   3.298522   .9100441     3.62   0.000     1.514868    5.082176
          x2 |   2.049418   .5098221     4.02   0.000     1.050185    3.048651
          x3 |   1.009372   .3101392     3.25   0.001     .4015108    1.617234
          x4 |   1.320595   .1769136     7.46   0.000     .9738509    1.667339
       _cons |   1.467998   8.093567     0.18   0.856    -14.3951     17.3311
------------------------------------------------------------------------------
```

```
Instrumented:  x1 x2
Instruments:   x3 x4 z1 z2 z3
```

Test the endogenous variable

```
estat endogenous x1

  Test of endogeneity (orthogonality conditions)
  Ho: variables are exogenous

  GMM C statistic chi2(1) =  13.3961  (p = 0.0003)

estat endogenous x2

  Test of endogeneity (orthogonality conditions)
  Ho: variables are exogenous

  GMM C statistic chi2(1) =  28.5438  (p = 0.0000)
```

Test the exogeneity of IV

```
estat overid

  Test of overidentifying restriction:

  Hansen's J chi2(1) = .014116 (p = 0.9054)
```

According to C statistic test, both x1 and x2 are endogenous variable. So, using GMM estimated result with z1 z2 z3 that have orthorgonal condition as IV would be appropriated result.

e. According to the estimated results of (d), give explanation of the differences between 2SLS and GMM estimated results in this case. (3 points)

In this case we have very similar results in 2SLS and GMM however the 2SLS have a little bit higher in R-square.

5. From the data set "`Midterm_q5_no.dta`":

a. Estimate the model assuming that the probability function is cumulative normal distribution function and logistic probability distribution. Can we compare the estimated coefficients of the two estimated functional forms? Why? or why not? Also, make interpret the estimated result of the Logit model (Overall test, individual test, pseudo *R2* , counted *R2*). (8 points)

cumulative normal distribution function

```
probit y x1 x2 x3

Iteration 0:   log likelihood = -112.90429
Iteration 1:   log likelihood =  -106.8431
Iteration 2:   log likelihood = -106.77169
Iteration 3:   log likelihood = -106.77166
Iteration 4:   log likelihood = -106.77166

Probit regression                               Number of obs     =        215
                                                LR chi2(3)        =      12.27
                                                Prob > chi2       =     0.0065
Log likelihood = -106.77166                     Pseudo R2         =     0.0543

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0032435   .0014378     2.26   0.024     .0004254    .0060615
          x2 |   .0104558   .0082123     1.27   0.203     -.00564     .0265515
          x3 |   .0913493   .0427701     2.14   0.033     .0075214    .1751773
       _cons |   .5058377   .1236092     4.09   0.000     .2635682    .7481072
------------------------------------------------------------------------------
```

logistic probability distribution

```
. logit y x1 x2 x3

Iteration 0:   log likelihood = -112.90429
Iteration 1:   log likelihood = -107.07485
Iteration 2:   log likelihood = -106.80412
Iteration 3:   log likelihood =  -106.8032
Iteration 4:   log likelihood =  -106.8032

Logistic regression                             Number of obs     =        215
                                                LR chi2(3)        =      12.20
                                                Prob > chi2       =     0.0067
Log likelihood =  -106.8032                     Pseudo R2         =     0.0540

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0056951   .0026061     2.19   0.029     .0005872    .0108029
          x2 |   .0201378   .0155022     1.30   0.194     -.010246    .0505216
          x3 |   .1668463   .0804444     2.07   0.038     .0091782    .3245144
       _cons |   .8143571   .2056236     3.96   0.000     .4113423    1.217372
------------------------------------------------------------------------------

. fitstat

Measures of Fit for logit of y

Log-Lik Intercept Only:     -112.904    Log-Lik Full Model:          -106.803
D(211):                      213.606    LR(3):                         12.202
                                        Prob > LR:                      0.007
McFadden's R2:                 0.054    McFadden's Adj R2:              0.019
Maximum Likelihood R2:         0.055    Cragg & Uhler's R2:             0.085
McKelvey and Zavoina's R2:     0.127    Efron's R2:                     0.050
Variance of y*:                3.766    Variance of error:              3.290
Count R2:                      0.781    Adj Count R2:                   0.000
AIC:                           1.031    AIC*n:                        221.606
BIC:                        -919.598    BIC':                           3.910
```

We cannot compare those 2 results coefficients since we use the different distribution

Interpretation of the estimated result of the Logit model

1.overall test we have LR chi2(3) = 12.20 which p-value < 0.05 so, we reject H0 which mean that the model is significant in overall test.

2.individual test we have insignificant at x2 coefficient

3.goodness of fit

Pseudo R2 = 0.0540 which don't have any interpretation we have to use It to compare with another model

Count R2: 0.781 mean the chance that model will have correct prediction and since it is higher than 0.5 it's considered as good model

b. Using logistic probability distribution, compute marginal effect at mean and at median. Make interpretation of marginal effects at mean of x_{1i}. (5 points)

marginal effect at mean

```
mfx

Marginal effects after logit
      y  = Pr(y) (predict)
         =  .80226499
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0009034       .0004    2.27   0.023   .000125  .001682   60.8585
      x2 |   .0031946      .00241    1.33   0.184  -.001521   .00791  -1.64096
      x3 |   .0264678      .01222    2.17   0.030   .002517  .050419   1.63388
------------------------------------------------------------------------------
```

marginal effect at median

```
. mfx, at(median)

Marginal effects after logit
      y  = Pr(y) (predict)
         =  .70198097
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0011914      .00059    2.01   0.045   .000027  .002356   .753216
      x2 |   .0042129      .00327    1.29   0.198  -.002205  .010631   .468091
      x3 |   .0349049      .01767    1.98   0.048   .000277  .069533    .17187
------------------------------------------------------------------------------
```

interpretation of marginal effects at mean of x_{1i}.

as liquidity ratio of firm i increase by 1 the probability that the firm i is dividend paying firm will increase by 0.0009034

(c) Perform hypothesis testing whether $\beta_1 = \beta_2 = \beta_3 = 0$ using LR-test and Wald test. Perform hypothesis testing whether $\beta_1 = \beta_2$ using LR-test. Make conclusion of the tests. (5 points)

Test B1=B2=B3 = 0 using LR-test

```
est store unres

. logit y

Iteration 0:   log likelihood = -112.90429
Iteration 1:   log likelihood = -112.90429

Logistic regression                             Number of obs     =        215
                                                LR chi2(0)        =       0.00
                                                Prob > chi2       =          .
Log likelihood = -112.90429                     Pseudo R2         =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   1.273816   .1650121     7.72   0.000     .9503987    1.597234
------------------------------------------------------------------------------

. est store res

. lrtest unres res

Likelihood-ratio test                                 LR chi2(3)  =     12.20
(Assumption: res nested in unres)                     Prob > chi2 =    0.0067
```

p-value < 0.05 reject H0
So, there is no evidence that B1,B2,B3 equal to 0

Test B1=B2=B3 = 0 using Waldtest

```
test x1 x2 x3

 ( 1)  [y]x1 = 0
 ( 2)  [y]x2 = 0
 ( 3)  [y]x3 = 0

           chi2(  3) =    9.61
         Prob > chi2 =    0.0222
```

p-value < 0.05 reject H0
So, there is no evidence that B1,B2,B3 equal to 0

Test B1=B2 using LR-test

```
constraint 1 x1=x2

. logit y x1 x2 x3, constraints(1)

Iteration 0:   log likelihood = -112.90429
Iteration 1:   log likelihood = -107.44788
Iteration 2:   log likelihood =  -107.2629
Iteration 3:   log likelihood =  -107.2625
Iteration 4:   log likelihood =  -107.2625

Logistic regression                             Number of obs     =        215
                                                Wald chi2(2)      =       9.44
Log likelihood =  -107.2625                     Prob > chi2       =     0.0089

 ( 1)  [y]x1 - [y]x2 = 0
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0055295   .0025378     2.18   0.029     .0005555    .0105034
          x2 |   .0055295   .0025378     2.18   0.029     .0005555    .0105034
          x3 |   .1442987   .0764271     1.89   0.059    -.0054956    .2940931
       _cons |   .8189633   .2048726     4.00   0.000     .4174204    1.220506
------------------------------------------------------------------------------

. est store constrained

. lrtest unres constrained

Likelihood-ratio test                                 LR chi2(1)  =      0.92
(Assumption: constrained nested in unres)             Prob > chi2 =    0.3378
```

p-value > 0.05 , Fail to reject H0
So, B1 and B2 are equaled

d. Why is threshold in computing Counted R^2 important? If we change the threshold, can the value of counted R^2 change? Why? (2 points)

It is important because it determines the result of decision. And if we change the threshold it possibly change the value of counted R square.

6. From the data set "`Midterm_q6_no.dta`":

a. Estimate model (13) using Panel Least Squares estimation method and PGLS assuming Heteroskedasticity, and test whether there exists Heteroskedasticity problem. What will happen if Heteroscedasticity occurs in the model (5 points)

PGLS assuming Heteroskedasticity

```
xtgls y x1 x2 x3, igls panels(heteroskedastic)
Iteration 1: tolerance = .00590371
Iteration 2: tolerance = .00177471
Iteration 3: tolerance = .00055434
Iteration 4: tolerance = .000177
Iteration 5: tolerance = .00005721
Iteration 6: tolerance = .00001862
Iteration 7: tolerance = 6.085e-06
Iteration 8: tolerance = 1.993e-06
Iteration 9: tolerance = 6.534e-07
Iteration 10: tolerance = 2.144e-07
Iteration 11: tolerance = 7.040e-08


Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        heteroskedastic
Correlation:   no autocorrelation

Estimated covariances      =       100          Number of obs     =      1,200
Estimated autocorrelations =         0          Number of groups  =        100
Estimated coefficients     =         4          Time periods      =         12
                                                Wald chi2(3)      =   33298.40
Log likelihood             = -6165.009          Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .3168738   .0100633    31.49   0.000     .2971501    .3365975
          x2 |   1.311977   .0139878    93.79   0.000     1.284562    1.339393
          x3 |  -.4630102    .011065   -41.84   0.000    -.4846972   -.4413232
       _cons |  -132.2058   6.140582   -21.53   0.000    -144.2411   -120.1705
------------------------------------------------------------------------------

. est store het
```

Panel Least Squares model

```
. xtgls y x1 x2 x3

Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        homoskedastic
Correlation:   no autocorrelation

Estimated covariances      =         1          Number of obs     =      1,200
Estimated autocorrelations =         0          Number of groups  =        100
Estimated coefficients     =         4          Time periods      =         12
                                                Wald chi2(3)      =   29882.41
Log likelihood             =  -6211.29          Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .3183087   .0109146    29.16   0.000     .2969164    .3397011
          x2 |   1.320714   .0149495    88.34   0.000     1.291413    1.350014
          x3 |  -.4642183    .011884   -39.06   0.000    -.4875104   -.4409262
       _cons |  -136.2145   6.645908   -20.50   0.000    -149.2402   -123.1887
------------------------------------------------------------------------------

. est store pgls

. local df=e(N_g)-1
```

test Heteroskedasticity

```
. lrtest het, df(`df')

Likelihood-ratio test                                 LR chi2(99) =     92.56
(Assumption: pgls nested in het)                      Prob > chi2 =    0.6629
```

From LR-test we fail to reject the null hypothesis which mean that the model has homoskedasticity.

If the model have heteroskedasticity it will cause high variance (not best) and t-test , F-test will be invalid however we can fix it with method like vce(robust)

b. Estimate the above three models including Panel Least Squares model (13), Fixed effects model (14), and Random-effects model (15). Perform fixed effects tests and random effects test, also state null hypothesis of the tests. Then, determine the most appropriated model. Also, give explanation of the choosing criterion (perform the tests), and make interpretation of the estimated models. (10 points)

Panel Least Squares model

```
xtgls y x1 x2 x3

Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        homoskedastic
Correlation:   no autocorrelation

Estimated covariances      =         1          Number of obs     =      1,200
Estimated autocorrelations =         0          Number of groups  =        100
Estimated coefficients     =         4          Time periods      =         12
                                                Wald chi2(3)      =   29882.41
Log likelihood             =  -6211.29          Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .3183087   .0109146    29.16   0.000     .2969164    .3397011
          x2 |   1.320714   .0149495    88.34   0.000     1.291413    1.350014
          x3 |  -.4642183    .011884   -39.06   0.000    -.4875104   -.4409262
       _cons |  -136.2145   6.645908   -20.50   0.000    -149.2402   -123.1887
------------------------------------------------------------------------------
```

Fixed effect model

```
. xtreg y x1 x2 x3 ,fe

Fixed-effects (within) regression               Number of obs     =      1,200
Group variable: id                              Number of groups  =        100

R-sq:                                           Obs per group:
     within  = 0.9502                                         min =         12
     between = 0.9848                                         avg =       12.0
     overall = 0.8846                                         max =         12

                                                F(3,1097)         =    6980.72
corr(u_i, Xb)  = 0.8057                         Prob > F          =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .1014426   .0044866    22.61   0.000     .0926393    .1102459
          x2 |   .6951028   .0085422    81.37   0.000      .678342    .7118637
          x3 |  -.4953168   .0041895  -118.23   0.000    -.5035372   -.4870965
       _cons |    208.659   4.373144    47.71   0.000     200.0784    217.2397
-------------+----------------------------------------------------------------
     sigma_u |  123.46108
     sigma_e |  14.376737
         rho |  .98662139   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Fixed effect test (H0: there is no specific fixed effect)

```
F test that all u_i=0: F(99, 1097) = 96.47                  Prob > F = 0.0000
```

From fixed effect test we can reject the H0 which mean that the fixed effect exists

```
. est store fe
```

random effect model

```
. xtreg y x1 x2 x3 ,re

Random-effects GLS regression                   Number of obs     =      1,200
Group variable: id                              Number of groups  =        100

R-sq:                                           Obs per group:
     within  = 0.8956                                         min =         12
     between = 0.9953                                         avg =       12.0
     overall = 0.9543                                         max =         12

                                                Wald chi2(3)      =    8707.50
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .2162513   .0090869    23.80   0.000     .1984414    .2340613
          x2 |   1.028607   .0152844    67.30   0.000     .9986503    1.058564
          x3 |  -.4779339   .0091412   -52.28   0.000    -.4958503   -.4600176
       _cons |   25.24251   8.000206     3.16   0.002     9.562392    40.92262
-------------+----------------------------------------------------------------
     sigma_u |  12.351151
     sigma_e |  14.376737
         rho |  .42464732   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. est store re
```

random effect test (H0: coefficients in FE = RE)

```
. hausman re fe

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       re           fe         Difference          S.E.
-------------+----------------------------------------------------------------
          x1 |    .2162513     .1014426        .1148087         .007902
          x2 |    1.028607     .6951028        .3335042        .0126745
          x3 |   -.4779339    -.4953168        .0173829        .0081246
------------------------------------------------------------------------------
                           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(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                         =     1451.21
               Prob>chi2 =      0.0000

.
```

From the random effect test we can reject the H0 which mean that fix effect is highly correlated with x so, we should use fix effect model

In conclusion, from the significant in both fixed effect and random effect test we can say that the fixed effect model is the most appropriated one in this case which fixed effects model is also significant in overall and individual test as well as high r-square.

c. What are the differences between Fixed effects estimation method and First difference estimation method? What are the differences between Fixed effects model and Random effects model? (3 points)

The difference between fixed effect and first difference is that first difference use the method difference-out between model at time t and model at time t-1 which require balance panel data but for the fixed effect we difference-out the cross-sectional group mean which can be used in the unbalance case as well.

For the random effects it's very similar to fixed effect but we difference-out not 1-1 or 100% of it but partly difference out because we think that the fixed effect is not 100% correlated with x .

d. Give explanation of Within R-squares, Overall R-squares, and Between R- squares of the estimated results of the Fixed-effects model. (2 points)

The between R square is the variance between seperate panel units that are accounted in the model
The within R square is variance within the panel units that are accounted in the model
and the R square overall is the average between them.