Answer Sheet Cover Page
Midterm Examination Semester 2/2020

(Readable handwriting and printed version are acceptable)

Student Name Phatakan Kanchanapradist

Student ID 6104641318

Course ID Course Title EE426

Lecturer Dr. Tatre Jantarakolica

Exam date Time

Total pages

***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 Phatakan Kanchanapradist

Date 21/3/2021

6104641318 Miss PHATAKAN KANCHANAPRADIST Midterm_q*_6.dta 06.zip

1.)

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

```
reg3 (y1 y2 x3) (y2 y1 x1 x2), ols


Multivariate regression
------------------------------------------------------------------------------
Equation              Obs    Parms        RMSE    "R-sq"     F-Stat        P
------------------------------------------------------------------------------
y1                     50        2    15.70603    0.8072      98.39   0.0000
y2                     50        3    6825.513    0.7149      38.45   0.0000
------------------------------------------------------------------------------


------------------------------------------------------------------------------
             |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0012299   .0002805     4.39   0.000      .000673    .0017869
          x3 |   .0064713   .0012305     5.26   0.000     .0040277    .0089148
       _cons |   49.03168   16.51628     2.97   0.004     16.23362    81.82974
-------------+----------------------------------------------------------------
y2           |
          y1 |   257.6683   34.31761     7.51   0.000     189.5203    325.8163
          x1 |   174.8604   95.66509     1.83   0.071    -15.11151    364.8323
          x2 |  -.0001503   .0002408    -0.62   0.534    -.0006285    .0003279
       _cons |  -23872.57   6722.813    -3.55   0.001    -37222.75   -10522.4
------------------------------------------------------------------------------
```

According to Simultaneous Equations Model, OLS estimators will be biased, inconsistent and inefficient.

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)

structural form :

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

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

reduced form :

$$y_{1t} = \Pi_{10} + \Pi_{11} x_{1t} + \Pi_{12} x_{2t} + \Pi_{13} x_{3t} + w_{1t}$$

$$y_{2t} = \Pi_{20} + \Pi_{21} x_{1t} + \Pi_{22} x_{2t} + \Pi_{23} x_{3t} + w_{2t}$$

Endo : y1 y2

Exo : x1 x2 x3

- estimate reduced form models using OLS and structural form models using IV technique from the predicted endogenous variables from reduced form estimated results.

```
. reg y1 x1 x2 x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     45.00
       Model |  44851.4833         3  14950.4944   Prob > F        =    0.0000
    Residual |  15283.4967        46  332.249929   R-squared       =    0.7458
-------------+----------------------------------   Adj R-squared   =    0.7293
       Total |    60134.98        49  1227.24449   Root MSE        =    18.228

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .4300947   .2497542     1.72   0.092    -.0726343    .9328238
          x2 |  -2.58e-07   6.42e-07    -0.40   0.689    -1.55e-06    1.03e-06
          x3 |   .0094901   .0011103     8.55   0.000     .0072553     .011725
       _cons |   20.70542   18.60798     1.11   0.272     -16.7505    58.16134
------------------------------------------------------------------------------

. predict y1hat, xb

. reg y2 x1 x2 x3
```

```
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     26.08
       Model |  4.7333e+09         3  1.5778e+09   Prob > F        =    0.0000
    Residual |  2.7829e+09        46  60498200.7   R-squared       =    0.6297
-------------+----------------------------------   Adj R-squared   =    0.6056
       Total |  7.5163e+09        49   153392984   Root MSE        =    7778.1

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   253.1533   106.5741     2.38   0.022     38.63086    467.6758
          x2 |  -.0001981    .000274    -0.72   0.473    -.0007497    .0003535
          x3 |   2.714804   .4737679     5.73   0.000     1.761159     3.66845
       _cons |  -21723.21    7940.32    -2.74   0.009    -37706.25   -5740.178
------------------------------------------------------------------------------

. predict y2hat, xb

. reg y1 y2hat x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(2, 47)        =     68.93
       Model |  44845.3725         2  22422.6863   Prob > F        =    0.0000
    Residual |  15289.6075        47  325.310797   R-squared       =    0.7457
-------------+----------------------------------   Adj R-squared   =    0.7349
       Total |    60134.98        49  1227.24449   Root MSE        =    18.036

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y2hat |   .0017233     .00096     1.80   0.079     -.000208    .0036545
          x3 |    .004819   .0033424     1.44   0.156    -.0019049     .011543
       _cons |    57.9742   25.06986     2.31   0.025     7.540145    108.4083
------------------------------------------------------------------------------

. reg y2 y1hat x1 x2
```

```
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     26.08
       Model |  4.7333e+09         3  1.5778e+09   Prob > F        =    0.0000
    Residual |  2.7829e+09        46    60498204   R-squared       =    0.6297
-------------+----------------------------------   Adj R-squared   =    0.6056
       Total |  7.5163e+09        49   153392984   Root MSE        =    7778.1

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y1hat |   286.0666   49.92227     5.73   0.000     185.5783    386.5549
          x1 |   130.1176   119.4766     1.09   0.282    -110.3763    370.6115
          x2 |  -.0001242   .0002759    -0.45   0.655    -.0006795    .0004311
       _cons |  -27646.34   8700.282    -3.18   0.003     -45159.1   -10133.58
------------------------------------------------------------------------------
```

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)

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

Three-stage least-squares regression

------------------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"       chi2        P
------------------------------------------------------------------------------
y1                    50       2    15.72072    0.7945     181.46   0.0000
y2                    50       3    6595.581    0.7106     108.87   0.0000
------------------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0017232   .0008367     2.06   0.039     .0000833    .0033632
          x3 |    .004819   .0029132     1.65   0.098    -.0008908    .0105289
       _cons |   57.97419   21.85119     2.65   0.008     15.14664    100.8017
```

```
-------------+----------------------------------------------------------------
y2           |
          y1 |   285.8013   42.29761     6.76   0.000     202.8995    368.7031
          x1 |   128.0824   100.4793     1.27   0.202    -68.85352    325.0182
          x2 |  -.0000953   .0001442    -0.66   0.509    -.0003778    .0001873
       _cons |  -27582.94   7366.319    -3.74   0.000    -42020.66   -13145.22
------------------------------------------------------------------------------

Endogenous variables:  y1 y2

Exogenous variables:   x1 x2 x3
```

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)

Single Equation

OLS : the estimated result will be biased , inconsistent and inefficient

2SLS : biased, consistent and Asymptotically efficient

System equations

3SLS : biased, consistent (can be inconsistent because of specification error), More Asymptotically efficient

Therefore, the advantage of system equation is that we get more Asymptotically efficient.

The disadvantage is when there exists specification error in one or many equations, the problem will spread through all equation in system.

2.)

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)

```
nl (lnC = {lnlamda}-(({beta}/{alpha})*(ln({theta}*(R^(-{alpha}))+(1-{theta})*W^(-{alpha})))),
init(lnlamda 1 theta 0.5 beta 0.5 alpha -0.5)

(obs = 252)


Iteration 0:  residual SS =  853.5392

Iteration 1:  residual SS =  281.2871

Iteration 2:  residual SS =  281.2867

Iteration 3:  residual SS =  281.2866
```

```
Iteration 4:  residual SS =  281.2866
Iteration 5:  residual SS =  281.2866
Iteration 6:  residual SS =  281.2866
Iteration 7:  residual SS =  281.2866
Iteration 8:  residual SS =  281.2866


      Source |       SS             df       MS
-------------+------------------------------------    Number of obs =        252
       Model |  45.117021           3   15.039007    R-squared     =     0.1382
    Residual |  281.28663         248  1.13422026    Adj R-squared =     0.1278
-------------+------------------------------------    Root MSE      =   1.064998
       Total |  326.40365         251  1.30041293    Res. dev.     =   742.8512

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    /lnlamda |   1.996386   .7811621     2.56   0.011     .4578285    3.534944
       /beta |   .8545114    .147163     5.81   0.000     .5646628     1.14436
      /alpha |   -.931979   .5400817    -1.73   0.086    -1.995711    .1317527
      /theta |   .2102829   .1833049     1.15   0.252    -.1507501    .5713159
------------------------------------------------------------------------------
  Parameter lnlamda taken as constant term in model & ANOVA table
est store lognls

. sca sigma=1/(1-(_b[/theta]))

. sca list sigma
     sigma =  1.2662762

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

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

```
       F(  3,   248) =   36.86
            Prob > F =    0.0000 < 0.05
Therefore , reject H0
```

b. 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)

```
nl (lnC = {lnlamda}-(({beta}/{alpha})*(ln({theta}*(R^(-{alpha}))+(1-{theta})*W^(-{alpha})))),
init(lnlamda 0.5 theta 0.1 beta 0.1 alpha -0.1)

(obs = 252)


Iteration 0:  residual SS =  4418.289
Iteration 1:  residual SS =  3876.752
Iteration 2:  residual SS =  500.3517
Iteration 3:  residual SS =  296.6092
Iteration 4:  residual SS =  290.3934
Iteration 5:  residual SS =   288.994
Iteration 6:  residual SS =  288.8673
Iteration 7:  residual SS =  288.8557
Iteration 8:  residual SS =  288.2211
Iteration 9:  residual SS =  284.7403
Iteration 10:  residual SS =  281.3093
Iteration 11:  residual SS =  281.2873
Iteration 12:  residual SS =  281.2867
Iteration 13:  residual SS =  281.2866
Iteration 14:  residual SS =  281.2866
Iteration 15:  residual SS =  281.2866
Iteration 16:  residual SS =  281.2866
Iteration 17:  residual SS =  281.2866
Iteration 18:  residual SS =  281.2866


      Source |       SS           df       MS
-------------+----------------------------------   Number of obs =       252
       Model |  45.117021          3   15.039007   R-squared     =    0.1382
    Residual |  281.28663        248  1.13422026   Adj R-squared =    0.1278
-------------+----------------------------------   Root MSE      =  1.064998
       Total |  326.40365        251  1.30041293   Res. dev.     =  742.8512
```

lnC	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
/lnlamda	1.996386	.7811621	2.56	0.011	.4578286	3.534944
/beta	.8545114	.147163	5.81	0.000	.5646628	1.14436
/alpha	-.9319789	.5400817	-1.73	0.086	-1.995711	.1317529
/theta	.2102829	.1833049	1.15	0.252	-.15075	.5713159

Parameter lnlamda taken as constant term in model & ANOVA table

Using different sets of initial values can give the different estimated results. However, ln this case, using log form can lead to more robust estimated > the estimated results become almost the same.

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)

```
nl (lnC = {lnlamda}-(({beta}/{alpha})*(ln({theta}*(R^(-{alpha}))+(1-{theta})*W^(-{alpha})))),
init(lnlamda 0.5 theta 0.1 beta 0.1 alpha -0.1)

(obs = 252)


Iteration 0:  residual SS =  4418.289
Iteration 1:  residual SS =  3876.752
Iteration 2:  residual SS =  500.3517
Iteration 3:  residual SS =  296.6092
Iteration 4:  residual SS =  290.3934
Iteration 5:  residual SS =   288.994
Iteration 6:  residual SS =  288.8673
Iteration 7:  residual SS =  288.8557
Iteration 8:  residual SS =  288.2211
Iteration 9:  residual SS =  284.7403
Iteration 10:  residual SS =  281.3093
Iteration 11:  residual SS =  281.2873
Iteration 12:  residual SS =  281.2867
Iteration 13:  residual SS =  281.2866
```

```
Iteration 14:  residual SS =  281.2866

Iteration 15:  residual SS =  281.2866

Iteration 16:  residual SS =  281.2866

Iteration 17:  residual SS =  281.2866

Iteration 18:  residual SS =  281.2866


      Source |       SS             df       MS
-------------+------------------------------------    Number of obs =        252
       Model |  45.117021           3   15.039007    R-squared     =     0.1382
    Residual |  281.28663         248  1.13422026    Adj R-squared =     0.1278
-------------+------------------------------------    Root MSE      =   1.064998
       Total |  326.40365         251  1.30041293    Res. dev.     =   742.8512

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    /lnlamda |   1.996386   .7811621     2.56   0.011     .4578286    3.534944
       /beta |   .8545114    .147163     5.81   0.000     .5646628     1.14436
      /alpha |  -.9319789   .5400817    -1.73   0.086    -1.995711    .1317529
      /theta |   .2102829   .1833049     1.15   0.252      -.15075    .5713159
------------------------------------------------------------------------------
  Parameter lnlamda taken as constant term in model & ANOVA table


. est store nls1


. nl (lnC = {lnlamda}-(({beta}/{alpha})*(ln({theta}*(R^(-{alpha}))+(1-{theta})*W^(-{alpha})))))),
init(lnlamda 0.5 theta 0.1 beta 0.1 alpha -0.1) eps(1e-1)

(obs = 252)


Iteration 0:  residual SS =  4418.289

Iteration 1:  residual SS =  3876.752

Iteration 2:  residual SS =  500.3517

Iteration 3:  residual SS =  296.6092

Iteration 4:  residual SS =  290.3934

Iteration 5:  residual SS =   288.994

Iteration 6:  residual SS =  288.8673
```

```
Iteration 7:  residual SS =  288.8557
Iteration 8:  residual SS =  288.2211
Iteration 9:  residual SS =  284.7403
Iteration 10:  residual SS =  281.3093
Iteration 11:  residual SS =  281.2873


      Source |       SS             df       MS
-------------+------------------------------------    Number of obs =        252
       Model |  45.116337           3   15.038779    R-squared     =     0.1382
    Residual |  281.28731         248  1.13422302    Adj R-squared =     0.1278
-------------+------------------------------------    Root MSE      =   1.064999
       Total |  326.40365         251  1.30041293    Res. dev.     =   742.8518

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    /lnlamda |   1.996213   .7820293     2.55   0.011     .4559475    3.536479
       /beta |   .8550992   .1471821     5.81   0.000     .5652128    1.144986
      /alpha |  -.9333344   .5161887    -1.81   0.072    -1.950007    .0833382
      /theta |    .208476   .1838451     1.13   0.258    -.1536208    .5705728
------------------------------------------------------------------------------
  Parameter beta taken as constant term in model & ANOVA table


. est store nls2


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

(obs = 252)


Iteration 0:  residual SS =  4418.289
Iteration 1:  residual SS =  3876.752
Iteration 2:  residual SS =  500.3517
Iteration 3:  residual SS =  296.6092
Iteration 4:  residual SS =  290.3934
Iteration 5:  residual SS =   288.994
Iteration 6:  residual SS =  288.8673
```

```
Iteration 7:  residual SS =  288.8557
Iteration 8:  residual SS =  288.2211
Iteration 9:  residual SS =  284.7403
Iteration 10:  residual SS =  281.3093
Iteration 11:  residual SS =  281.2873
Iteration 12:  residual SS =  281.2867
Iteration 13:  residual SS =  281.2866
Iteration 14:  residual SS =  281.2866
Iteration 15:  residual SS =  281.2866
Iteration 16:  residual SS =  281.2866
Iteration 17:  residual SS =  281.2866
Iteration 18:  residual SS =  281.2866
Iteration 19:  residual SS =  281.2866
Iteration 20:  residual SS =  281.2866
Iteration 21:  residual SS =  281.2866
Iteration 22:  residual SS =  281.2866
Iteration 23:  residual SS =  281.2866
Iteration 24:  residual SS =  281.2866
Iteration 25:  residual SS =  281.2866
Iteration 26:  residual SS =  281.2866
Iteration 27:  residual SS =  281.2866
Iteration 28:  residual SS =  281.2866
Iteration 29:  residual SS =  281.2866


      Source |       SS            df       MS
-------------+----------------------------------    Number of obs =        252
       Model |  11625.479           4  2906.36982    R-squared     =     0.9764
    Residual |  281.28663         248  1.13422026    Adj R-squared =     0.9760
-------------+----------------------------------    Root MSE      =   1.064998
       Total |  11906.766         252  47.2490711    Res. dev.     =   742.8512

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    /lnlamda |   1.996386   .7811622     2.56   0.011     .4578281    3.534944
```

```
       /beta |   .8545116    .147163     5.81   0.000      .564663     1.14436
      /alpha |  -.9319801   .5400797    -1.73   0.086    -1.995708    .1317477
      /theta |   .2102824   .1833053     1.15   0.252    -.1507512    .5713161
------------------------------------------------------------------------------
```

```
. est store nls3
```

```
. est table nls1 nls2 nls3, star(.1 .05 .01) stat(N rss r2 r2_a)
```

```
----------------------------------------------------------
    Variable |     nls1            nls2            nls3
-------------+--------------------------------------------
lnlamda      |
       _cons |  1.9963865**     1.9962133**     1.9963862**
-------------+--------------------------------------------
beta         |
       _cons |  .85451141***    .85509917***    .85451156***
-------------+--------------------------------------------
alpha        |
       _cons | -.93197895*     -.93333443*     -.93198015*
-------------+--------------------------------------------
theta        |
       _cons |  .21028292       .20847602       .21028244
-------------+--------------------------------------------
Statistics   |
           N |        252             252             252
         rss |  281.28663       281.28731       281.28663
          r2 |  .13822462       .13822253        .9763759
        r2_a |  .12779992        .1277978       .97599487
----------------------------------------------------------
                          legend: * p<.1; ** p<.05; *** p<.01
```

According to the above estimated results, it indicated that using log-form can lead to more robust estimated results.

d. Why do we prefer to estimate nonlinear regression model in log-form instead of its original functional form?

Because if we use log-form, it can speed up the process and we can get robust estimated results.

3.)

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

```
. program ml_logit
  1.    args lnf theta
  2.    quietly replace `lnf'=ln(1/(1+exp(-`theta'))) if $ML_y1==1
  3.    quietly replace `lnf'=ln(1-(1/(1+exp(-`theta')))) if $ML_y1==0
  4. end

.
end of do-file

. 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
------------------------------------------------------------------------------

. est store newton

.  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
------------------------------------------------------------------------------

. est store bhhh

. 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
------------------------------------------------------------------------------

. est store bfgs

. est table newton bhhh bfgs, star(.1 .05 .01) stat(N ll chi2 p)

-------------------------------------------------------------
    Variable |    newton           bhhh            bfgs
-------------+-----------------------------------------------
          x1 |  .25087707**     .25088813**     .25088266**
          x2 | -.56265628***   -.56265912***   -.56265833***
       _cons |  .62497609***    .62495691***    .62497069***
-------------+-----------------------------------------------
           N |        400             400             400
          ll | -227.67158      -227.67158      -227.67158
        chi2 |  66.430535       81.09243       66.430831
           p |  3.757e-15      2.460e-18       3.756e-15
-------------------------------------------------------------
                         legend: * p<.1; ** p<.05; *** p<.01
```

The estimated results using different algorithm can be different because of the different computation function.

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

```
. test x1 x2
```

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

           chi2(  2) =   66.43
         Prob > chi2 =    0.0000

. est store ur

. . ml model lf ml_logit (y=)

. 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 =  -275.0498
Iteration 2:   log likelihood =  -275.0498

                                                Number of obs     =        400
                                                Wald chi2(0)      =          .
Log likelihood =  -275.0498                     Prob > chi2       =          .

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |  -.2107769   .1005559    -2.10   0.036    -.4078627   -.0136911
------------------------------------------------------------------------------

. est store res

. lrtest ur res

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

Make comparison between the two tests. Which test is preferable? Why? (5 points)

```
Both tests > Ho is rejected

LR is prefer since it is optimal power.
```

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)

no we cannot because in MLE, F test is not reliable.

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)

```
. program ml_probit_het
  1.    args lnf theta sigma
  2.    tempvar z s
  3.    quietly g double `s'=exp(`sigma')
  4.    quietly g double `z'=`theta'/`s'
  5.    quietly replace `lnf'=ln(normal(`z')) if $ML_y1==1
  6.    quietly replace `lnf'=ln(1-normal(`z')) if $ML_y1==0
  7. end
end of do-file


.  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 = -259.02903  (not concave)
Iteration 2:   log likelihood = -254.52067
Iteration 3:   log likelihood = -234.56196
Iteration 4:   log likelihood = -229.81568
Iteration 5:   log likelihood = -225.79455
Iteration 6:   log likelihood = -225.58385
Iteration 7:   log likelihood = -225.58236
```

```
Iteration 8:   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    .7125442
------------------------------------------------------------------------------


. est store hetprob


. lrtest hetprob ur

Likelihood-ratio test                                 LR chi2(1)  =      4.18
(Assumption: ur nested in hetprob)                    Prob > chi2 =    0.0409


Prob > chi2 =    0.0409 < 0.05
Reject H0 > there is significant heteroskedasticity.
```

- No we can not use LM or Wald since they require only one model.

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)

> it depends on the distribution of the regression model; the calculated parameters are not always t-distributed in the MLE case. Consequently, the Z-test is be used to conduct individual experiments.

> because we use log-likelihood function, so we have log-likelihood value instead.

> we can use log-likelihood value to compare between two estimated results.

5.)

(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 R 2 , counted R 2). (8 points)

> Can we compare the estimated coefficients of the two estimated functional forms? Why? or why not?

We cannot compare the estimated coefficients of the two models. Since they are not nested model.

```
. probit y x1 x2 x3

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -113.64516
Iteration 2:   log likelihood = -113.47712
Iteration 3:   log likelihood =  -113.4769
Iteration 4:   log likelihood =  -113.4769

Probit regression                               Number of obs     =        215
                                                LR chi2(3)        =      21.73
                                                Prob > chi2       =     0.0001
Log likelihood =  -113.4769                     Pseudo R2         =     0.0874

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0045328   .0014479     3.13   0.002     .0016949    .0073706
          x2 |   .0104545   .0082353     1.27   0.204    -.0056863    .0265953
          x3 |   .1107471   .0426902     2.59   0.009     .0270759    .1944184
       _cons |    .273115   .1193808     2.29   0.022     .0391328    .5070972
------------------------------------------------------------------------------

. logit y x1 x2 x3

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -114.09273
Iteration 2:   log likelihood = -113.53546
Iteration 3:   log likelihood = -113.53345
```

```
Iteration 4:   log likelihood = -113.53345

Logistic regression                             Number of obs     =        215
                                                LR chi2(3)        =      21.62
                                                Prob > chi2       =     0.0001
Log likelihood = -113.53345                     Pseudo R2         =     0.0869

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0078182   .0026109     2.99   0.003     .0027009    .0129354
          x2 |   .0200053   .0155107     1.29   0.197    -.0103952    .0504058
          x3 |    .198448    .079765     2.49   0.013     .0421114    .3547845
       _cons |   .4322113   .1937858     2.23   0.026     .0523981    .8120244
------------------------------------------------------------------------------

. fitstat

Measures of Fit for logit of y

Log-Lik Intercept Only:       -124.343   Log-Lik Full Model:          -113.533
D(211):                        227.067   LR(3):                         21.620
                                         Prob > LR:                      0.000
McFadden's R2:                   0.087   McFadden's Adj R2:              0.055
ML (Cox-Snell) R2:               0.096   Cragg-Uhler(Nagelkerke) R2:     0.140
McKelvey & Zavoina's R2:         0.190   Efron's R2:                     0.088
Variance of y*:                  4.061   Variance of error:              3.290
Count R2:                        0.735   Adj Count R2:                   0.000
AIC:                             1.093   AIC*n:                        235.067
BIC:                          -906.138   BIC':                          -5.508
BIC used by Stata:             248.549   AIC used by Stata:            235.067
```

Overall test `LR chi2(3) = 21.73`

individual test : Z-test

pseudo R 2 : 1-(-113.53345/-124.343)

counted R 2 : `Count R2: 0.735`

(b) Using logistic probability distribution, compute marginal effect at mean and at median. Make interpretation of marginal effects at mean of x1

```
. mfx

Marginal effects after logit
      y  = Pr(y) (predict)
         =  .76813757
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0013924      .00044    3.17   0.002   .000531  .002254   60.8706
      x2 |    .003563      .00271    1.32   0.188  -.001741  .008867  -1.61707
      x3 |    .035344      .01351    2.62   0.009   .008857  .061831   1.62292
------------------------------------------------------------------------------

. mfx, at(median)

Marginal effects after logit
      y  = Pr(y) (predict)
         =  .61720905
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0018471      .00066    2.82   0.005   .000561  .003133   .780464
      x2 |   .0047265      .00368    1.28   0.199  -.002489  .011942   .468091
      x3 |   .0468857       .0194    2.42   0.016   .008857  .084915   .151382
------------------------------------------------------------------------------
```

Make interpretation of marginal effects at mean of x1

- If x1 chage by 1 unit (from 60.8706 to 61.8706), P(y=1) will increase by 0.0013924, holding other x at their mean.

(c) Perform hypothesis testing

```
. logit y x1 x2 x3
```

```
Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -114.09273
Iteration 2:   log likelihood = -113.53546
Iteration 3:   log likelihood = -113.53345
Iteration 4:   log likelihood = -113.53345

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =      21.62
                                                  Prob > chi2       =     0.0001
Log likelihood = -113.53345                       Pseudo R2         =     0.0869

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0078182   .0026109     2.99   0.003     .0027009    .0129354
          x2 |   .0200053   .0155107     1.29   0.197    -.0103952    .0504058
          x3 |    .198448    .079765     2.49   0.013     .0421114    .3547845
       _cons |   .4322113   .1937858     2.23   0.026     .0523981    .8120244
------------------------------------------------------------------------------

. test (x1=0) (x2=0) (x3=0)

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

           chi2(  3) =   16.26
         Prob > chi2 =    0.0010

. constraint 1 x1=x2

. logit y x1 x2 x3,constraints(1)

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -114.29289
```

```
Iteration 2:   log likelihood = -113.86391
Iteration 3:   log likelihood = -113.86201
Iteration 4:   log likelihood = -113.86201

Logistic regression                             Number of obs     =        215
                                                Wald chi2(2)      =      16.50
Log likelihood = -113.86201                     Prob > chi2       =     0.0003

 ( 1)  [y]x1 - [y]x2 = 0

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |    .007643   .0025485     3.00   0.003     .0026479     .012638
          x2 |    .007643   .0025485     3.00   0.003     .0026479     .012638
          x3 |   .1792934   .0756342     2.37   0.018     .0310532    .3275337
       _cons |   .4379841   .1932953     2.27   0.023     .0591323    .8168359
------------------------------------------------------------------------------

. est store con

. logit y x1 x2 x3

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -114.09273
Iteration 2:   log likelihood = -113.53546
Iteration 3:   log likelihood = -113.53345
Iteration 4:   log likelihood = -113.53345

Logistic regression                             Number of obs     =        215
                                                LR chi2(3)        =      21.62
                                                Prob > chi2       =     0.0001
Log likelihood = -113.53345                     Pseudo R2         =     0.0869

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
          x1 |   .0078182   .0026109     2.99   0.003     .0027009    .0129354
          x2 |   .0200053   .0155107     1.29   0.197    -.0103952    .0504058
          x3 |    .198448    .079765     2.49   0.013     .0421114    .3547845
       _cons |   .4322113   .1937858     2.23   0.026     .0523981    .8120244
------------------------------------------------------------------------------


. est store full


. lrtest full con

Likelihood-ratio test                                 LR chi2(1)  =      0.66
(Assumption: con nested in full)                      Prob > chi2 =    0.4176
```

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)

- Threshold can change the value of counted R^2 because it can changes the number of correct prediction which is used in calculating counted r^2

4.

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).

```
. tsset time
        time variable:  time, 1 to 1326
                delta:  1 unit


. g dr=f.r-r
(1 missing value generated)


. 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) 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
```

```
numerical derivatives are approximate
flat or discontinuous region encountered
Iteration 0:   GMM criterion Q(b) =  .00001195
Iteration 1:   GMM criterion Q(b) =  8.825e-06  (backed up)
Iteration 2:   GMM criterion Q(b) =  6.165e-06  (not concave)
Iteration 3:   GMM criterion Q(b) =  3.790e-06  (backed up)
Iteration 4:   GMM criterion Q(b) =  3.098e-06
Iteration 5:   GMM criterion Q(b) =  9.848e-07
Iteration 6:   GMM criterion Q(b) =  1.867e-09
Iteration 7:   GMM criterion Q(b) =  4.232e-13

Step 2
Iteration 0:   GMM criterion Q(b) =  2.771e-09
Iteration 1:   GMM criterion Q(b) =  3.178e-16

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 |  -.0023857   .0011649    -2.05   0.041    -.0046688   -.0001026
       /beta |   .0004306   .0002873     1.50   0.134    -.0001326    .0009938
     /sigma2 |   .0005163   .0003278     1.57   0.115    -.0001263    .0011589
      /gamma |   .0932382   .1800569     0.52   0.605    -.2596669    .4461433
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
```

```
Instruments for equation 4: _cons


. est store Unrestricted


. gmm (dr-{alpha}) ((dr-{alpha})*r) ((dr-{alpha})^2-{sigma2}) (((dr-{alpha})^2- {sigma2})*r)
winitia
> l(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) =  .00001195
Iteration 1:   GMM criterion Q(b) =  4.124e-08
Iteration 2:   GMM criterion Q(b) =  4.124e-08


Step 2
Iteration 0:   GMM criterion Q(b) =  .00772783
Iteration 1:   GMM criterion Q(b) =  .00546904
Iteration 2:   GMM criterion Q(b) =  .00546904


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 |  -.0008144   .0006886    -1.18   0.237    -.0021641    .0005353
     /sigma2 |   .0004396    .000293     1.50   0.134    -.0001347    .0010139
------------------------------------------------------------------------------
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) = 7.24648 (p = 0.0267)


. est store Merton


. gmm (dr-{alpha}-{beta}*r) ((dr-{alpha}-{beta}*r)*r) ((dr-{alpha}-{beta}*r)^2- {sigma2}) (((dr-
{alp

> ha}-{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) =  .00001195

Iteration 1:   GMM criterion Q(b) =  3.109e-10

Iteration 2:   GMM criterion Q(b) =  3.044e-10


Step 2

Iteration 0:   GMM criterion Q(b) =  .00057191

Iteration 1:   GMM criterion Q(b) =  .00019007

Iteration 2:   GMM criterion Q(b) =  .00019007


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 |  -.0027001   .0009815    -2.75   0.006    -.0046239   -.0007763
       /beta |   .0005339   .0002002     2.67   0.008     .0001415    .0009263
     /sigma2 |   .0005981   .0003003     1.99   0.046     9.57e-06    .0011867
------------------------------------------------------------------------------
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) = .251838 (p = 0.6158)

. est store Vasicek

. est table  Unrestricted Merton Varsicek, star(0.1 0.05 0.01) stat(N J)
estimation result Varsicek not found
r(111);

. est table  Unrestricted Merton Vasicek, star(0.1 0.05 0.01) stat(N J)

----------------------------------------------------------
    Variable | Unrestricted       Merton          Vasicek
-------------+--------------------------------------------
alpha        |
       _cons | -.00238571**    -.00081437      -.00270009***
-------------+--------------------------------------------
beta         |
       _cons |  .00043061                       .00053387***
-------------+--------------------------------------------
```

```
sigma2       |
       _cons |  .00051631       .00043961       .00059814**
-------------+----------------------------------------------
gamma        |
       _cons |  .09323815
-------------+----------------------------------------------
Statistics   |
           N |       1325            1325            1325
           J |  4.211e-13       7.2464785       .25183811
------------------------------------------------------------
                           legend: * p<.1; ** p<.05; *** p<.01


. test (_b[/gamma]=0)
equation gamma not found
r(111);


. est restore Unrestricted
(results Unrestricted are active now)


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


 ( 1)  [gamma]_cons = 0


           chi2(  1) =    0.27
         Prob > chi2 =    0.6046
```

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

Vasick is the most appropriated model because of wald test (accept H0) and also accept HO from J test (This Model can be trusted).

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

Based on, some E (x,u) not equal to zero, so it can cause biased

d. Estimate the model (10) using GMM and 2SLS by employing z1i, z2i, and z3i as instrumental variables for x1i and x2i. Perform the test to check whether GMM is appropriated.

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

```
Instrumental variables (GMM) regression          Number of obs   =        500
                                                 Wald chi2(4)    =     162.88
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.6085
GMM weight matrix: Robust                        Root MSE        =     20.051

------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   3.122913   .9709225     3.22   0.001      1.21994    5.025886
          x2 |   2.154339   .5222639     4.13   0.000     1.130721    3.177958
          x3 |   1.051275   .3117477     3.37   0.001     .4402609     1.66229
          x4 |   1.305762   .1765515     7.40   0.000     .9597271    1.651796
       _cons |   1.036629   7.417985     0.14   0.889    -13.50236    15.57561
------------------------------------------------------------------------------
Instrumented:  x1 x2
Instruments:   x3 x4 z1 z2 z3

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

Instrumental variables (2SLS) regression         Number of obs   =        500
                                                 Wald chi2(4)    =     151.08
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.6085
                                                 Root MSE        =     20.052

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   3.125296   .9506451     3.29   0.001     1.262066    4.988526
          x2 |   2.152402   .5201568     4.14   0.000     1.132913    3.171891
          x3 |   1.049297   .3106005     3.38   0.001     .4405316    1.658063
          x4 |    1.30648   .1795184     7.28   0.000     .9546299    1.658329
       _cons |   1.061581   8.036584     0.13   0.895    -14.68983      16.813
------------------------------------------------------------------------------
```

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

```
. estat endogenous x1 x2

  Tests of endogeneity
  Ho: variables are exogenous

  Durbin (score) chi2(2)          =  458.015  (p = 0.0000)
  Wu-Hausman F(2,493)             =  2689.05  (p = 0.0000)
```

p-value < 0.05 > Ho is rejected. Therefore, x1 and x2 are endogenous variables. So, GMM estimated result is the most appropriated estimated result.

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

The different is in 2sls we use x1hat x2hat while in GMM, we use z1 z2 z3

6.

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)

```
. xtgls y x1 x2 x3, igls panels(heteroskedastic) nolog


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

.  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

. lrtest het, df(`df')

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

With LR-chi2 test, 0.6629>0.05, the null hypothesis of no heteroscedasticity is not rejected, thus, there exists insignificant heteroscedasticity problem.

What will happen if Heteroscedasticity occurs in the model?

The estimators become just LUE, it is not the most. However, this is not that serious problem.

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)

```
. 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
------------------------------------------------------------------------------

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

. est store fe

.  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

. hausman fe re

                ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       fe           re         Difference          S.E.
-------------+----------------------------------------------------------------
          x1 |    .1014426     .2162513       -.1148087               .
          x2 |    .6951028     1.028607       -.3335042               .
          x3 |   -.4953168    -.4779339       -.0173829               .
------------------------------------------------------------------------------
```

```
                           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    chi2<0 ==> model fitted on these
                                        data fails to meet the asymptotic
                                        assumptions of the Hausman test;
                                        see suest for a generalized test

. 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
According to the fixed effects test, there is the significant fixed effects. According to the
significant Hausman test (p-value of 0.0000<0.05), the null hypothesis of the test is rejected,
thus, the fixed effects model is more appropriated than random effects model.
```

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)

> First difference can use with balance panels only, while Fixed effects can use for

unbalanced panels, not just balanced panels.

- FE model is subset of RE. In FE, we assume that alpha was correlated with x. but In RE, the correlated can high or not high depends on lambda. Ex. If lambda =1, this is also FE

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

From FE model, if we would like to compare, we can use Overall R^2 because we calculate from actual value. Moreover, within R^2 is fine too but between R^2 is not.