Mai Surintraboon
9_6204640178

Question 1

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)

```
. tsset id
        time variable:  id, 1 to 50
                delta:  1 unit


. reg3 (y1 y2 x3)(y2 y1 x1 x2), ols
Multivariate regression
------------------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"     F-Stat        P
------------------------------------------------------------------------------
y1                    50       2    16.00066    0.8129     102.09   0.0000
y2                    50       3     9499.16    0.6639      30.29   0.0000
------------------------------------------------------------------------------
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0009514   .0001947     4.89   0.000     .0005647    .0013381
          x3 |   .0072546   .0010926     6.64   0.000      .005085    .0094243
       _cons |   46.76431   16.79595     2.78   0.006     13.41088    80.11773
-------------+----------------------------------------------------------------
y2           |
          y1 |    302.627   47.22599     6.41   0.000     208.8455    396.4084
          x1 |   239.0346   133.1835     1.79   0.076    -25.44145    503.5106
          x2 |  -.0004611   .0003289    -1.40   0.164    -.0011142     .000192
       _cons |  -36348.02   9096.518    -4.00   0.000     -54411.9   -18284.14
------------------------------------------------------------------------------
```

The consequences of using OLS in this case is that the nature of the model has simultaneous bias as the dependent variable (y1 and y2) are function of each other. Thus, the estimated result are biased, inconsistent, and inefficient.

Mai Surintraboon
9_6204640178

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)

Estimate the model using 2SLS

```
. reg3 ( y1 y2 x3 )( y2 y1 x1 x2 ), 2sls inst( x1 x2 x3 )


Two-stage least-squares regression

------------------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"     F-Stat        P
------------------------------------------------------------------------------
y1                    50       2     16.4114    0.8032      88.52   0.0000
y2                    50       3     9502.61    0.6636      24.21   0.0000
------------------------------------------------------------------------------


------------------------------------------------------------------------------
             |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0012558   .0005291     2.37   0.020     .0002052    .0023065
          x3 |   .0061047   .0021635     2.82   0.006     .0018085    .0104009
       _cons |    54.4282   21.18721     2.57   0.012     12.35459    96.50181
-------------+----------------------------------------------------------------
y2           |
          y1 |   293.9937   61.45979     4.78   0.000     171.9467    416.0407
          x1 |   253.5942   148.8156     1.70   0.092    -41.92397    549.1125
          x2 |  -.0004718   .0003326    -1.42   0.159    -.0011322    .0001887
       _cons |  -35239.13   10406.85    -3.39   0.001    -55905.06   -14573.19
------------------------------------------------------------------------------
Endogenous variables:  y1 y2
Exogenous variables:   x1 x2 x3
------------------------------------------------------------------------------
```

Mai Surintraboon
9_6204640178

State reduced form and structural form of these simultaneous equation models

Structural Form

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

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

Reduced Form

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

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

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

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

$$\boxed{y_{1t} = \pi_0 + \pi_1 x_{1t} + \pi_2 x_{2t} + \pi_3 x_{3t} + w_{1t}} \quad \text{——①'}$$

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

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

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

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

$$\boxed{y_{2t} = \pi_4 + \pi_5 x_{1t} + \pi_6 x_{2t} + \pi_7 x_{3t} + w_{2t}} \quad \text{——②'}$$

Where

$$\pi_0 = \frac{\beta_{10} + \gamma_{12}\beta_{20}}{(1 - \gamma_{12}\gamma_{21})} \qquad \pi_4 = \frac{(\beta_{20} + \gamma_{21}\beta_{10})}{(1 - \gamma_{21}\gamma_{12})}$$

$$\pi_1 = \frac{\gamma_{12}\beta_{12}}{(1 - \gamma_{12}\gamma_{21})} \qquad \pi_5 = \frac{\beta_{21}}{(1 - \gamma_{21}\gamma_{12})}$$

$$\pi_2 = \frac{\gamma_{12}\beta_{22}}{(1 - \gamma_{12}\gamma_{21})} \qquad \pi_6 = \frac{\beta_{22}}{(1 - \gamma_{21}\gamma_{12})}$$

$$\pi_3 = \frac{\beta_{13}}{(1 - \gamma_{12}\gamma_{21})} \qquad \pi_7 = \frac{\gamma_{21}\beta_{13}}{(1 - \gamma_{21}\gamma_{12})}$$

$$w_{1t} = \frac{\gamma_{12} u_{2t} + u_{1t}}{(1 - \gamma_{12}\gamma_{21})} \qquad w_{2t} = \frac{(\gamma_{21} u_{1t} + u_{2t})}{(1 - \gamma_{21}\gamma_{12})}$$

Specify endogenous variables and exogenous variables

Endogenous variables: y1 y2

Exogenous variables: x1 x2 x3

Mai Surintraboon
9_6204640178

Estimate reduced form models using OLS

```
. reg y1 x1 x2 x3
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     44.24
       Model |  47757.2453         3  15919.0818   Prob > F        =    0.0000
    Residual |  16552.5347        46  359.837712   R-squared       =    0.7426
-------------+----------------------------------   Adj R-squared   =    0.7258
       Total |    64309.78        49  1312.44449   Root MSE        =    18.969

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .5371558   .2555819     2.10   0.041      .022696    1.051616
          x2 |  -5.62e-07   6.52e-07    -0.86   0.393    -1.87e-06    7.50e-07
          x3 |    .009485   .0011637     8.15   0.000     .0071426    .0118274
       _cons |   16.03845   19.32441     0.83   0.411    -22.85958    54.93647
------------------------------------------------------------------------------

. predict y1hat,xb

. reg y2 x1 x2 x3
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     17.37
       Model |  6.5593e+09         3  2.1864e+09   Prob > F        =    0.0000
    Residual |  5.7898e+09        46   125865910   R-squared       =    0.5312
-------------+----------------------------------   Adj R-squared   =    0.5006
       Total |  1.2349e+10        49   252022302   Root MSE        =     11219

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   411.5147   151.1579     2.72   0.009     107.2496    715.7797
          x2 |  -.0006369   .0003854    -1.65   0.105    -.0014126    .0001389
          x3 |   2.788536   .6882408     4.05   0.000     1.403179    4.173892
       _cons |  -30523.92   11428.97    -2.67   0.010    -53529.24   -7518.605
------------------------------------------------------------------------------

. predict y2hat,xb
```

Mai Surintraboon
9_6204640178

Estimate structural form models using IV technique from the predicted endogenous variables from reduced form estimated results

```
. reg y1 y2hat x3
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(2, 47)        =     67.39
       Model |   47682.413         2  23841.2065   Prob > F        =    0.0000
    Residual |   16627.367        47  353.773766   R-squared       =    0.7414
-------------+----------------------------------   Adj R-squared   =    0.7304
       Total |    64309.78        49  1312.44449   Root MSE        =    18.809

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y2hat |   .0012558   .0006064     2.07   0.044      .000036    .0024757
          x3 |   .0061047   .0024795     2.46   0.018     .0011166    .0110929
       _cons |   54.42821   24.28236     2.24   0.030      5.57839     103.278
------------------------------------------------------------------------------

. reg y2 y1hat x1 x2
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     17.37
       Model |  6.5593e+09         3  2.1864e+09   Prob > F        =    0.0000
    Residual |  5.7898e+09        46   125865907   R-squared       =    0.5312
-------------+----------------------------------   Adj R-squared   =    0.5006
       Total |  1.2349e+10        49   252022302   Root MSE        =     11219

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y1hat |   293.9937   72.56083     4.05   0.000     147.9364    440.0511
          x1 |   253.5942    175.695     1.44   0.156    -100.0615      607.25
          x2 |  -.0004718   .0003926    -1.20   0.236    -.0012621    .0003186
       _cons |  -35239.13   12286.56    -2.87   0.006    -59970.69   -10507.57
------------------------------------------------------------------------------
```

Mai Surintraboon
9_6204640178

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 ), 3sls inst( x1 x2 x3 )
Three-stage least-squares regression
------------------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"       chi2        P
------------------------------------------------------------------------------
y1                    50       2    15.91145    0.8032     188.34   0.0000
y2                    50       3    9131.701    0.6624      79.22   0.0000
------------------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0012558   .0005129     2.45   0.014     .0002505    .0022612
          x3 |   .0061047   .0020976     2.91   0.004     .0019936    .0102158
       _cons |    54.4282   20.54177     2.65   0.008     14.16708    94.68932
-------------+----------------------------------------------------------------
y2           |
          y1 |   288.9321   58.21035     4.96   0.000     174.8419    403.0223
          x1 |    264.349   141.3615     1.87   0.061    -12.71442    541.4124
          x2 |  -.0003806   .0002713    -1.40   0.161    -.0009124    .0001512
       _cons |  -35180.32    9981.31    -3.52   0.000    -54743.33   -15617.31
------------------------------------------------------------------------------
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).

The differences among the three model is that…

Mai Surintraboon
9_6204640178

1.OLS does not take in account the simultaneous bias that occur because of the nature of the equations. If we run OLS, we will obtain bias, inconsistent, and inefficient estimators.

2.2SLS takes the simultaneous bias that occur into account. This method starts with predicting the endogenous variable from the reduced form model and then go on to use the predicted value as IV. After that it uses these IVs to estimate the structural form of the model. The estimated results will still be bias, but now it is consistent and asymptotically efficient.

3.3SLS takes the simultaneous bias that occur into account. This method has the same procedure as 2SLS but adding another procedure: FGLS after 2SLS. FGLS is perform by running the structural form again as a whole system. The estimated results obtained from 3SLS will be bias, consistent, and more asymptotically efficient than other methods. More efficiency is the advantage of running the system equation instead of the single equations.

However, we need to be careful because if there exist any specification error in any equations. The error will spread throughout the whole system. This is the disadvantage of running system equation instead of single equation like OLS or 2SLS.

Question 2

a. Estimate the model (4) using NLS estimation method using initial values of $\ln\lambda=1$, $\theta=0.5$, $\beta=0.5$, $\alpha=-0.5$. Determine the estimated value of efficiency parameter (λ), distribution parameter (θ), parameter (β), and substitution parameter (α), and elasticity of substitution (σ).

```
. *a

. nl (lnC = {lnlambda}-(({beta}/{alpha})*(ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha})))))), in

> it(lnlambda 1 theta 0.5 beta 0.5 alpha -0.5)

(obs = 250)


Iteration 0:  residual SS =  320.6977

Iteration 1:  residual SS =  301.8529

Iteration 2:  residual SS =  301.6507

Iteration 3:  residual SS =  301.6506

Iteration 4:  residual SS =  301.6506

Iteration 5:  residual SS =  301.6506

Iteration 6:  residual SS =  301.6506


      Source |       SS           df       MS                    
-------------+----------------------------------    Number of obs =       250
       Model |  54.484926          3  18.1616418    R-squared     =    0.1530
    Residual |  301.65058        246  1.22622188    Adj R-squared =    0.1427
-------------+----------------------------------    Root MSE      =  1.107349
       Total |  356.13551        249  1.43026308    Res. dev.     =  756.4214
```

```
------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   /lnlambda |   1.920144   .7400844     2.59   0.010     .4624331    3.377854
        /beta |   .8625099   .1355566     6.36   0.000     .5955103    1.129509
       /alpha |  -.8267275   .4666301    -1.77   0.078    -1.745827    .0923724
       /theta |   .2624028   .1786237     1.47   0.143    -.0894242    .6142297
------------------------------------------------------------------------------
  Parameter lnlambda taken as constant term in model & ANOVA table


. est store nls1


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


. sca list sigma
     sigma =  1.3557534
```

b. What will happen if we change initial values to ln□=0.5, □=0.1, □=0.1, □= –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)

```
. *b
. nl (lnC = {lnlambda}-(({beta}/{alpha})*(ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha})))), in
> it(lnlambda 0.5 theta 0.1 beta 0.1 alpha -0.1)
(obs = 250)
Iteration 0:  residual SS =  4392.861
Iteration 1:  residual SS =  3727.718
Iteration 2:  residual SS =  357.4753
Iteration 3:  residual SS =  348.7848
Iteration 4:  residual SS =  324.1562
Iteration 5:  residual SS =  324.0813
Iteration 6:  residual SS =  323.8134
Iteration 7:  residual SS =  323.4937
Iteration 8:  residual SS =  323.1047
Iteration 9:  residual SS =  322.4562
Iteration 10:  residual SS =  321.2419
Iteration 11:  residual SS =  318.1517
Iteration 12:  residual SS =   314.016
```

```
Iteration 13:  residual SS =  301.7794
Iteration 14:  residual SS =  301.6506
Iteration 15:  residual SS =  301.6506
Iteration 16:  residual SS =  301.6506
Iteration 17:  residual SS =  301.6506

      Source |       SS           df       MS
-------------+----------------------------------    Number of obs =       250
       Model |  54.484926          3  18.1616418    R-squared     =    0.1530
    Residual |  301.65058        246  1.22622188    Adj R-squared =    0.1427
-------------+----------------------------------    Root MSE      =  1.107349
       Total |  356.13551        249  1.43026308    Res. dev.     =  756.4214

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   /lnlambda |   1.920144   .7400844     2.59   0.010     .4624332    3.377854
       /beta |   .8625098   .1355565     6.36   0.000     .5955103    1.129509
      /alpha |  -.8267272   .4666279    -1.77   0.078    -1.745823    .0923684
      /theta |   .2624029   .1786237     1.47   0.143    -.0894241    .6142298
------------------------------------------------------------------------------
  Parameter lnlambda taken as constant term in model & ANOVA table


. est store nls2


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


. sca list sigma
     sigma =  1.3557537
```

The estimated result from (a) and (b) are very similar, but they are a little different if we carefully look at the estimated parameters value on the far right of the number. Moreover, the number of iterations are different as in(a) it is 6 times and (b) is 17 times. This difference occurs because (a) and (b) have different initial values.

Mai Surintraboon
9_6204640178

c. 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)

(i)0.1 or (1e-1)

```
. *c.1

. nl (lnC = {lnlambda}-(({beta}/{alpha})*(ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha})))), in

> it(lnlambda 0.5 theta 0.1 beta 0.1 alpha -0.1)eps(1e-1)

(obs = 250)


Iteration 0:  residual SS =  4392.861

Iteration 1:  residual SS =  3727.718

Iteration 2:  residual SS =  357.4753

Iteration 3:  residual SS =  348.7848

Iteration 4:  residual SS =  324.1562

Iteration 5:  residual SS =  324.0813

Iteration 6:  residual SS =  323.8134

Iteration 7:  residual SS =  323.4937

Iteration 8:  residual SS =  323.1047

Iteration 9:  residual SS =  322.4562

Iteration 10:  residual SS =  321.2419

Iteration 11:  residual SS =  318.1517

Iteration 12:  residual SS =   314.016

Iteration 13:  residual SS =  301.7794

Iteration 14:  residual SS =  301.6506


      Source |       SS           df       MS
-------------+----------------------------------    Number of obs =       250
       Model |  54.484891          3  18.1616304    R-squared     =    0.1530
    Residual |  301.65062        246  1.22622202    Adj R-squared =    0.1427
-------------+----------------------------------    Root MSE      =  1.107349
       Total |  356.13551        249  1.43026308    Res. dev.     =  756.4214

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
```

```
-------------+------------------------------------------------------------------
   /lnlambda |    1.92033   .7402206     2.59   0.010     .4623511    3.378308
       /beta |   .8624709   .1355743     6.36   0.000     .5954365    1.129505
      /alpha |  -.8255423   .4737031    -1.74   0.083    -1.758574    .1074889
      /theta |   .2626096   .1824575     1.44   0.151    -.0967685    .6219877
------------------------------------------------------------------------------
  Parameter beta taken as constant term in model & ANOVA table


. est store nls3


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


. sca list sigma
     sigma =  1.3561338
```

ii) (1e-15) with maximum iteration of 40

```
. *c.2

. nl (lnC = {lnlambda}-(({beta}/{alpha})*(ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha})))), in

> it(lnlambda 0.5 theta 0.1 beta 0.1 alpha -0.1)eps(1e-15) iter(40)

(obs = 250)


Iteration 0:  residual SS =  4392.861
Iteration 1:  residual SS =  3727.718
Iteration 2:  residual SS =  357.4753
Iteration 3:  residual SS =  348.7848
Iteration 4:  residual SS =  324.1562
Iteration 5:  residual SS =  324.0813
Iteration 6:  residual SS =  323.8134
Iteration 7:  residual SS =  323.4937
Iteration 8:  residual SS =  323.1047
Iteration 9:  residual SS =  322.4562
Iteration 10:  residual SS =  321.2419
Iteration 11:  residual SS =  318.1517
Iteration 12:  residual SS =   314.016
Iteration 13:  residual SS =  301.7794
Iteration 14:  residual SS =  301.6506
Iteration 15:  residual SS =  301.6506
Iteration 16:  residual SS =  301.6506
```

```
Iteration 17:  residual SS =  301.6506
Iteration 18:  residual SS =  301.6506
Iteration 19:  residual SS =  301.6506
Iteration 20:  residual SS =  301.6506


      Source |       SS           df       MS                      
-------------+----------------------------------   Number of obs =       250
       Model |  11449.247          4  2862.31181   R-squared     =    0.9743
    Residual |  301.65058        246  1.22622188   Adj R-squared =    0.9739
-------------+----------------------------------   Root MSE      =  1.107349
       Total |  11750.898        250  47.0035913   Res. dev.     =  756.4214

------------------------------------------------------------------------------
         lnC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   /lnlambda |   1.920144   .7400844     2.59   0.010     .4624331    3.377854
       /beta |   .8625099   .1355566     6.36   0.000     .5955103    1.129509
      /alpha |  -.8267275   .4666305    -1.77   0.078    -1.745828    .0923732
      /theta |   .2624027   .1786237     1.47   0.143    -.0894242    .6142297
------------------------------------------------------------------------------

. est store nls4

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

. sca list sigma
     sigma =  1.355753
.
end of do-file
```

Mai Surintraboon
9_6204640178

Making comparison between new results and (b)

```
. est table nls2 nls3 nls4, star(.1 .05 .01) stat(N rss r2 r2_a)


----------------------------------------------------------
    Variable |     nls2            nls3            nls4
-------------+--------------------------------------------
lnlambda     |
       _cons |  1.9201436**     1.9203297**     1.9201435**
-------------+--------------------------------------------
beta         |
       _cons |  .86250983***    .86247092***    .86250986***
-------------+--------------------------------------------
alpha        |
       _cons | -.82672716*     -.82554234*     -.82672751*
-------------+--------------------------------------------
theta        |
       _cons |  .26240287       .26260962       .26240274
-------------+--------------------------------------------
Statistics   |
           N |        250             250             250
         rss |  301.65058       301.65062       301.65058
          r2 |  .15298931       .15298921       .97432957
        r2_a |  .14265991       .14265981       .97391217
----------------------------------------------------------
                          legend: * p<.1; ** p<.05; *** p<.01
```

The estimated result in (b, (c)(i), and (c)(ii) are different because they have different initial values, convergence value, and max iteration. (b) and (c) have different initial value. (b), (c)(i), and (c)(ii) have different convergence value. Moreover, in (c)(ii), we have set out maximum iteration at 40.

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

We prefer log-form because using it can lead to more robust estimated results. The estimated results are more likely to be significant.

Mai Surintraboon
9_6204640178

Question 3

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)

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

. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood = -272.63079
rescale:       log likelihood = -271.87577
Iteration 0:   log likelihood = -271.87577
Iteration 1:   log likelihood = -229.84968
Iteration 2:   log likelihood = -229.48827
Iteration 3:   log likelihood = -229.48797
Iteration 4:   log likelihood = -229.48797

                                                Number of obs     =        400
                                                Wald chi2(2)      =      61.78
Log likelihood = -229.48797                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .5449328   .1230867     4.43   0.000     .3036873    .7861782
          x2 |  -.5155686   .0715727    -7.20   0.000    -.6558485   -.3752887
       _cons |   .0717861   .1992166     0.36   0.719    -.3186713    .4622436
------------------------------------------------------------------------------

. est store logit1

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

. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood = -272.63079
rescale:       log likelihood = -271.87577
```

```
Iteration 0:   log likelihood = -271.87577
Iteration 1:   log likelihood = -229.85671
Iteration 2:   log likelihood =  -229.4898
Iteration 3:   log likelihood = -229.48797
Iteration 4:   log likelihood = -229.48797

                                                Number of obs     =        400
                                                Wald chi2(2)      =      62.12
Log likelihood = -229.48797                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |                 OPG
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .5449256   .1221639     4.46   0.000     .3054887    .7843625
          x2 |  -.5155846   .0723328    -7.13   0.000    -.6573542    -.373815
       _cons |   .0718396   .2050023     0.35   0.726    -.3299575    .4736366
------------------------------------------------------------------------------

. est store logit2

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

. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood = -272.63079
rescale:       log likelihood = -271.87577
Iteration 0:   log likelihood = -271.87577
Iteration 1:   log likelihood = -250.87722  (backed up)
Iteration 2:   log likelihood = -230.14003  (backed up)
Iteration 3:   log likelihood = -229.92959
Iteration 4:   log likelihood = -229.48942
Iteration 5:   log likelihood =   -229.488
Iteration 6:   log likelihood = -229.48797
Iteration 7:   log likelihood = -229.48797

                                                Number of obs     =        400
```

```
                                                Wald chi2(2)      =      61.78
Log likelihood = -229.48797                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .5449327   .1230867     4.43   0.000     .3036873    .7861782
          x2 |  -.5155686   .0715727    -7.20   0.000    -.6558484   -.3752887
       _cons |   .0717861   .1992166     0.36   0.719    -.3186714    .4622435
------------------------------------------------------------------------------
```

```
. est store logit3
```

```
. est table logit1 logit2 logit3, star(.1 .05 .01)stat(N ll chi2 p)
```

```
-------------------------------------------------------------
    Variable |    logit1          logit2          logit3
-------------+-----------------------------------------------
          x1 |  .54493275***    .54492561***    .54493275***
          x2 |  -.5155686***   -.51558458***   -.51556859***
       _cons |  .07178611       .07183955       .07178607
-------------+-----------------------------------------------
           N |        400             400             400
          ll | -229.48797      -229.48797      -229.48797
        chi2 |  61.779312       62.124104        61.77931
           p |  3.844e-14       3.235e-14       3.844e-14
-------------------------------------------------------------
                           legend: * p<.1; ** p<.05; *** p<.01
```

The 3 estimated results are different because the models employ different algorithm, which lead to the usage of different computation function.

b. Perform hypothesis testing whether beta1=beta2=0 using LR-test and Wald test. Make comparison between the two tests. Which test is preferable? Why? (5 points)

In (b), we use the data from the estimated models using MLE with Newton-Ralphson algorithm.

LR Test

Estimate Unrestricted Model

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

```
. ml maximize
```

Mai Surintraboon
9_6204640178

```
initial:       log likelihood = -277.25887
alternative:   log likelihood = -272.63079
rescale:       log likelihood = -271.87577
Iteration 0:   log likelihood = -271.87577
Iteration 1:   log likelihood = -229.84968
Iteration 2:   log likelihood = -229.48827
Iteration 3:   log likelihood = -229.48797
Iteration 4:   log likelihood = -229.48797

                                                Number of obs     =        400
                                                Wald chi2(2)      =      61.78
Log likelihood = -229.48797                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .5449328   .1230867     4.43   0.000     .3036873    .7861782
          x2 |  -.5155686   .0715727    -7.20   0.000    -.6558485   -.3752887
       _cons |   .0717861   .1992166     0.36   0.719    -.3186713    .4622436
------------------------------------------------------------------------------
. est store logit1
```

Estimate Restricted Model

```
. ml model lf ml_logit (y=)
. ml maximize

initial:       log likelihood = -277.25887
alternative:   log likelihood = -272.63079
rescale:       log likelihood = -271.87577
Iteration 0:   log likelihood = -271.87577
Iteration 1:   log likelihood = -271.45072
Iteration 2:   log likelihood =  -271.4507

                                                Number of obs     =        400
                                                Wald chi2(0)      =          .
Log likelihood =  -271.4507                     Prob > chi2       =          .

------------------------------------------------------------------------------
```

```
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |  -.3433333   .1014771    -3.38   0.001    -.5422247   -.1444419
------------------------------------------------------------------------------

. est store logit1_res
```

Perform LR Test

```
. lrtest logit1 logit1_res

Likelihood-ratio test                                 LR chi2(2)  =     83.93
(Assumption: logit1_res nested in logit1)             Prob > chi2 =    0.0000
```

Wald Test

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

           chi2(  2) =   61.78
         Prob > chi2 =    0.0000

. ml model lf ml_logit (y=)
. ml maximize
initial:       log likelihood = -277.25887
alternative:   log likelihood = -272.63079
rescale:       log likelihood = -271.87577
Iteration 0:   log likelihood = -271.87577
Iteration 1:   log likelihood = -271.45072
Iteration 2:   log likelihood =  -271.4507

                                                Number of obs     =        400
                                                Wald chi2(0)      =          .
Log likelihood =  -271.4507                     Prob > chi2       =          .

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |  -.3433333   .1014771    -3.38   0.001    -.5422247   -.1444419
------------------------------------------------------------------------------
```

Mai Surintraboon
9_6204640178

The different between the LR and Wald test is that the LR test uses two models: unrestricted and restricted, but the Wald test uses only the unrestricted model. The LR test is preferable because it has the optimal power. However, if our restricted model is complicated, the Wald test should be employed.

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)

The reason is that MLE and OLS (which uses F-test) employ different estimation method. MLE aims to maximize the likelihood function whereas OLS aims to minimize the sum of square residuals. The F-test is constructed out of the minimizing least square concept; thus we cannot employ F-test as overall test in MLE.

$$\text{F-test} = \frac{(\text{RSS of restricted model} - \text{RSS of unrestricted model})/(k-1)}{(\text{RSS of unrestricted model}/(n-k)}$$

d. Estimate the models with heteroskedasticity using MLE with Newton-Ralphson algorithm. Perform LR-test whether there exists significant heteroskedasticity

```
. doedit "C:\Users\user\Documents\BE TU\Year2\EE426\MIDTERM EXAM\Midterm_q3 - ml_probit.do"

. do "C:\Users\user\AppData\Local\Temp\STD02000000.tmp"

. program ml_probit

  1.    args lnf theta

  2.    tempvar z

  3.    quietly g double `z'=`theta'

  4.    quietly replace `lnf'=ln(normal(`z')) if $ML_y1==1

  5.    quietly replace `lnf'=ln(1-normal(`z')) if $ML_y1==0

  6. end

end of do-file

. ml model lf ml_probit ( y= x1 x2 )

. ml maximize

initial:       log likelihood = -277.25887

alternative:   log likelihood = -281.53481

rescale:       log likelihood = -271.60653

Iteration 0:   log likelihood = -271.60653

Iteration 1:   log likelihood = -229.86554

Iteration 2:   log likelihood = -229.59367

Iteration 3:   log likelihood =  -229.5935

Iteration 4:   log likelihood =  -229.5935


                                                Number of obs     =        400

                                                Wald chi2(2)      =      69.64

Log likelihood =  -229.5935                     Prob > chi2       =     0.0000
```

Mai Surintraboon
9_6204640178

```
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .3244745   .0723427     4.49   0.000     .1826855    .4662635
          x2 |  -.3082745   .0408367    -7.55   0.000     -.388313   -.2282361
       _cons |   .0469691   .1205919     0.39   0.697    -.1893866    .2833248
------------------------------------------------------------------------------

. est store probit

. doedit "C:\Users\user\Documents\BE TU\Year2\EE426\MIDTERM EXAM\Midterm_q3 -
ml_probit_het.do"

. do "C:\Users\user\AppData\Local\Temp\STD02000000.tmp"

. 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 = -281.46023

rescale:       log likelihood = -271.58192

rescale eq:    log likelihood = -271.36437

Iteration 0:   log likelihood = -271.36437

Iteration 1:   log likelihood = -236.90428

Iteration 2:   log likelihood = -228.17359

Iteration 3:   log likelihood = -227.35327

Iteration 4:   log likelihood =  -227.3515

Iteration 5:   log likelihood =  -227.3515


                                                Number of obs     =        400

                                                Wald chi2(2)      =      66.09

Log likelihood =  -227.3515                     Prob > chi2       =     0.0000


------------------------------------------------------------------------------
```

```
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
eq1          |
          x1 |   .2970344    .072702     4.09   0.000     .1545411    .4395277
          x2 |   -.315029   .0411731    -7.65   0.000    -.3957267   -.2343312
       _cons |   .0907444   .1164622     0.78   0.436    -.1375174    .3190061
-------------+----------------------------------------------------------------
eq2          |
          x3 |  -.2484118   .1178887    -2.11   0.035    -.4794693   -.0173543
------------------------------------------------------------------------------
```

```
. est store hetprob

. lrtest hetprob probit

Likelihood-ratio test                                 LR chi2(1)  =      4.48
(Assumption: probit nested in hetprob)                Prob > chi2 =    0.0342
```

From the above LR test, since the p-value is less than 0.05, the null hypothesis of no heteroskedasticity can be rejected. Thus, there is significant heteroskedasticity in this model.

d.(continue)In this case, can we perform Wald-test or LM-test to test heteroskedasticity? Why? or why not? (5 points)

We can perform LM-test to test heteroskedasticity because it takes 2 models into consideration. We can compare between the model that we suspect of having heteroskedasticity and the one we are sure that it does not have.

We cannot use Wald-test as it only takes the unrestricted model into account. In this case, we need to compare between the 2 model to be able to perform this test.

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)

We employ z-test instead of t-test in MLE because we assume asymptotic normal to be able to conduct the test.

Moreover, we do not have R-square in MLE because R-square is created out of the notion that we try to minimize the least square (OLS) not the notion of maximizing the log-likelihood function in the MLE.

R-square = 1- (Sum of square residual/Total sum of square)

Mai Surintraboon
9_6204640178

Question 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)

```
. *Unresticted

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

> a})) (((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) =  .00001194

Iteration 1:   GMM criterion Q(b) =  8.809e-06  (backed up)

Iteration 2:   GMM criterion Q(b) =  6.154e-06  (not concave)

Iteration 3:   GMM criterion Q(b) =  4.182e-06  (backed up)

Iteration 4:   GMM criterion Q(b) =  3.231e-06

Iteration 5:   GMM criterion Q(b) =  7.789e-08

Iteration 6:   GMM criterion Q(b) =  3.974e-08


Step 2

Iteration 0:   GMM criterion Q(b) =  .00006118

Iteration 1:   GMM criterion Q(b) =  1.335e-09

Iteration 2:   GMM criterion Q(b) =  7.626e-17


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 |  -.0023874   .0011648    -2.05   0.040    -.0046703   -.0001046
       /beta |    .000431   .0002874     1.50   0.134    -.0001322    .0009942
     /sigma2 |   .0005156    .000328     1.57   0.116    -.0001272    .0011585
      /gamma |   .0933678   .1802519     0.52   0.604    -.2599195    .4466551
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
Instruments for equation 4: _cons

. est store Unrestricted

. *Merton
. gmm (dr-{alpha}) ((dr-{alpha})*r) ((dr-{alpha})^2-{sigma2}) (((dr-{alpha})^2-
{sigma2})*r) winiti
> al(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) =  .00001194
Iteration 1:   GMM criterion Q(b) =  4.130e-08
Iteration 2:   GMM criterion Q(b) =  4.129e-08

Step 2
Iteration 0:   GMM criterion Q(b) =  .00777293
Iteration 1:   GMM criterion Q(b) =  .00548447
Iteration 2:   GMM criterion Q(b) =  .00548447

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 |  -.0008163   .0006884    -1.19   0.236    -.0021655    .0005329
     /sigma2 |   .0004387   .0002931     1.50   0.134    -.0001357    .0010131
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
Instruments for equation 4: _cons
. est store Merton

. *Vasicek
. gmm (dr-{alpha}-{beta}*r) ((dr-{alpha}-{beta}*r)*r) ((dr-{alpha}-{beta}*r)^2-
{sigma2}) (((dr-{al
> pha}-{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) =  .00001194
Iteration 1:   GMM criterion Q(b) =  3.110e-10
Iteration 2:   GMM criterion Q(b) =  3.045e-10

Step 2
Iteration 0:   GMM criterion Q(b) =  .00057424
Iteration 1:   GMM criterion Q(b) =  .00019027
Iteration 2:   GMM criterion Q(b) =  .00019027

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 |  -.0027025   .0009807    -2.76   0.006    -.0046247   -.0007803
       /beta |   .0005344   .0002001     2.67   0.008     .0001422    .0009266
     /sigma2 |   .0005976   .0003004     1.99   0.047     8.90e-06    .0011863
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
Instruments for equation 4: _cons
. est store Vasicek
```

Evaluation of the Merton model

Looking at the individual z-test of Merton's estimated result, it could be seen that they are all insignificant at 95% confidence level. The p-value of the estimated are more than 0.05; thus, the null hypothesis cannot be rejected.

Moreover, when performing the Hansen's j-test for Merton, the result is as follows…

```
. estat overid

  Test of overidentifying restriction:

  Hansen's J chi2(2) = 7.26693 (p = 0.0264)
```

The p-value are less than 0.05, meaning that we can reject the null hypothesis that over-identified moment conditions are satisfied. Thus, there is no need to do GMM in this case.

Therefore, Merton is not an appropriate model to use.

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

```
. est table Unrestricted Merton Vasicek, star(0.1 0.05 0.01) stat(N J)
-------------------------------------------------------------
    Variable | Unrestricted       Merton          Vasicek
-------------+-----------------------------------------------
alpha        |
       _cons | -.00238743**    -.00081632      -.00270248***
-------------+-----------------------------------------------
beta         |
       _cons |  .00043104                       .0005344***
-------------+-----------------------------------------------
sigma2       |
```

```
       _cons |  .00051564       .00043867       .00059761**
-------------+----------------------------------------------
gamma        |
       _cons |  .09336781
-------------+----------------------------------------------
Statistics   |
           N |       1325            1325            1325
           J |  1.010e-13       7.2669287       .25210658
------------------------------------------------------------
                           legend: * p<.1; ** p<.05; *** p<.01
```

***J-test for Vasicek**

```
estat overid
  Test of overidentifying restriction:
  Hansen's J chi2(1) = .252107 (p = 0.6156)
```

***Wald Test**

```
. est restore Unrestricted
(results Unrestricted are active now)
```

. ***Test Unrestricted vs Merton**

```
. test (_b[/beta]=0) (_b[/gamma]=0)
 ( 1)  [beta]_cons = 0
 ( 2)  [gamma]_cons = 0
           chi2(  2) =    7.80
         Prob > chi2 =    0.0203
```

. ***Test Unrestricted vs Vasicek**

```
. test (_b[/gamma]=0)
 ( 1)  [gamma]_cons = 0
           chi2(  1) =    0.27
         Prob > chi2 =    0.6045
```

The most appropriated model is Vasicek as can be shown in the individual significant test (z-test), Hansen's J-test, and the Wald test above (these are the selected criteria).

Vasicek individual estimated results are all significant at 95% confidence level. Also, it passes the J-test. The p-value is 0.6156, meaning that we cannot reject the null hypothesis that over-identified moment conditions are satisfied.

Moreover, when we employ the Wald test, Vasicek p-value is 0.6045, meaning that we cannot reject the null hypothesis that gamma=0, which is exactly the condition specify in the Vasicek model.

Mai Surintraboon
9_6204640178

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)

The estimated results would be bias because there exists correlation between the error term and the independent variables: X1 and X2. The zero conditional mean assumption is violated.

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. (5 points)

GMM

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

Instrumental variables (GMM) regression          Number of obs   =        500
                                                 Wald chi2(4)    =     177.10
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.6164
GMM weight matrix: Robust                        Root MSE        =     19.826

------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   3.240971   .9464757     3.42   0.001     1.385913    5.096029
          x2 |   2.108794   .5198782     4.06   0.000     1.089851    3.127736
          x3 |   .9878937   .3079089     3.21   0.001     .3844034    1.591384
          x4 |   1.402216      .1718     8.16   0.000     1.065494    1.738937
       _cons |   1.188526   7.361258     0.16   0.872    -13.23927    15.61633
------------------------------------------------------------------------------
Instrumented:  x1 x2
Instruments:   x3 x4 z1 z2 z3
```

2SLS

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

Instrumental variables (2SLS) regression         Number of obs   =        500
                                                 Wald chi2(4)    =     166.27
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.6167
                                                 Root MSE        =     19.818
```

```
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   3.236032   .9276263     3.49   0.000     1.417918    5.054147
          x2 |   2.115172   .5216801     4.05   0.000     1.092697    3.137646
          x3 |   .9922347   .3065636     3.24   0.001      .391381    1.593088
          x4 |   1.400136   .1771847     7.90   0.000      1.05286    1.747411
       _cons |    1.08386   7.980978     0.14   0.892    -14.55857    16.72629
------------------------------------------------------------------------------
```

Instrumented: x1 x2

Instruments: x3 x4 z1 z2 z3

Test to check whether GMM is appropriated

```
. estat overid

  Test of overidentifying restriction:

  Hansen's J chi2(1) = .078075 (p = 0.7799)
```

To determine appropriateness, I will use the individual Z-test and the J-test as the criteria. As can be seen in the individual Z-test of GMM's estimated results above, the results are all significant as the estimates' p-value are all less than 0.05.

Moreover, from the J-test, the p-value is 0.7799. Thus, we cannot reject the null hypothesis that the over-identified moment conditions are satisfied.

All in all, the two test results show that GMM is appropriated.

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

The difference between 2SLS and GMM estimated results stems from the usage of different standard errors. GMM employs the robust S.E., while 2sls employs the regular S.E. If there exist any autocorrelation or heteroskedasticity, the usage of robust S.E. could ameliorate the problem.

Question 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-R 2, counted R 2) (8 points)

Probability function: cumulative normal distribution

```
. probit y x1 x2 x3

Iteration 0:   log likelihood = -124.34324

Iteration 1:   log likelihood = -113.46552

Iteration 2:   log likelihood = -113.28881
```

```
Iteration 3:   log likelihood = -113.28856
Iteration 4:   log likelihood = -113.28856

Probit regression                                 Number of obs     =        215
                                                  LR chi2(3)        =      22.11
                                                  Prob > chi2       =     0.0001
Log likelihood = -113.28856                       Pseudo R2         =     0.0889

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0045465   .0014497     3.14   0.002     .0017051    .0073879
          x2 |   .0108073   .0082528     1.31   0.190    -.0053678    .0269824
          x3 |   .1137361   .0429917     2.65   0.008     .0294739    .1979983
       _cons |   .2682669   .1196918     2.24   0.025     .0336753    .5028584
------------------------------------------------------------------------------
. fitstat
Measures of Fit for probit of y
Log-Lik Intercept Only:     -124.343     Log-Lik Full Model:         -113.289
D(211):                      226.577     LR(3):                       22.109
                                         Prob > LR:                    0.000
McFadden's R2:                 0.089     McFadden's Adj R2:            0.057
Maximum Likelihood R2:         0.098     Cragg & Uhler's R2:           0.143
McKelvey and Zavoina's R2:     0.204     Efron's R2:                   0.090
Variance of y*:                1.256     Variance of error:            1.000
Count R2:                      0.735     Adj Count R2:                 0.000
AIC:                           1.091     AIC*n:                      234.577
BIC:                        -906.628     BIC':                        -5.997
```

Probability function: logistic distribution

```
. logit y x1 x2 x3
Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -113.91256
Iteration 2:   log likelihood =  -113.3249
Iteration 3:   log likelihood = -113.32272
Iteration 4:   log likelihood = -113.32272

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =      22.04
                                                  Prob > chi2       =     0.0001
```

Mai Surintraboon
9_6204640178

```
Log likelihood = -113.32272                     Pseudo R2         =     0.0886

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0078609   .0026177     3.00   0.003     .0027303    .0129916
          x2 |   .0207881   .0155648     1.34   0.182    -.0097184    .0512945
          x3 |   .2048633   .0806633     2.54   0.011     .0467662    .3629603
       _cons |   .4230449   .1942924     2.18   0.029     .0422387     .803851
------------------------------------------------------------------------------
. fitstat
Measures of Fit for logit of y
Log-Lik Intercept Only:     -124.343     Log-Lik Full Model:         -113.323
D(211):                      226.645     LR(3):                       22.041
                                         Prob > LR:                    0.000
McFadden's R2:                 0.089     McFadden's Adj R2:            0.056
Maximum Likelihood R2:         0.097     Cragg & Uhler's R2:           0.142
McKelvey and Zavoina's R2:     0.195     Efron's R2:                   0.091
Variance of y*:                4.085     Variance of error:            3.290
Count R2:                      0.735     Adj Count R2:                 0.000
AIC:                           1.091     AIC*n:                      234.645
BIC:                        -906.559     BIC':                        -5.929

. estat clas
Logistic model for y
              -------- True --------
Classified |         D            ~D  |      Total
-----------+--------------------------+-----------
     +     |       158            57  |        215
     -     |         0             0  |          0
-----------+--------------------------+-----------
   Total   |       158            57  |        215

Classified + if predicted Pr(D) >= .5
True D defined as y != 0
--------------------------------------------------
Sensitivity                     Pr( +| D)  100.00%
Specificity                     Pr( -|~D)    0.00%
```

```
Positive predictive value       Pr( D| +)   73.49%
Negative predictive value       Pr(~D| -)       .%
--------------------------------------------------
False + rate for true ~D        Pr( +|~D)  100.00%
False - rate for true D         Pr( -| D)    0.00%
False + rate for classified +   Pr(~D| +)   26.51%
False - rate for classified -   Pr( D| -)       .%
--------------------------------------------------
Correctly classified                        73.49%
--------------------------------------------------
```

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

No, because they employ different probability distribution function. Logit employs logistic distribution, while Probit uses cumulative normal.

Make interpret the estimated result of the Logit model (Overall test, individual test, pseudo-R 2, counted R 2)

From the LR-Chi square test, the p-value is 0.0001. Therefore, we can reject the null hypothesis that the independent variables are not significant. Logit is overall significant.

However, from the individual Z-test, not all independent variables are significant. X1 and X3 are significant, but X2 is not as it has the p-value=0.182 (>0.05).

Logit has the value of pseudo-R2 equals to 0.0886. This R2 cannot be used to make interpretation. It can only be used for comparison.

Moreover, Logit’s overall counted R2 equals to 73.49%. However, the counted R2 when y=1 is 100.00% and when y=0 is 0.00%. This is not a good sign because these three numbers should be very similar to each other and more than 50% if we have a good estimate.

Thus, we could state that it is not best to use Logit model in this case.

b. Using logistic probability distribution, compute marginal effect at mean and at median. Make interpretation of marginal effects at mean of x1i. (5 points)

```
. mfx

Marginal effects after logit
      y  = Pr(y) (predict)
         =  .76907563
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0013961      .00044    3.18   0.001   .000536  .002256   60.8702
      x2 |   .0036919       .0027    1.37   0.172  -.001607  .008991  -1.59279
      x3 |   .0363834      .01358    2.68   0.007    .00977  .062997   1.63362
------------------------------------------------------------------------------
```

Mai Surintraboon
9_6204640178

Make interpretation of marginal effects at mean of x1i

If X1i(liquidity ratio of firm i) increases by 1 unit, the predicted probability that the listed company in the stock market is a dividend paying firm will increase by 0.13961%, by holding the log of firm i size and profitability index of firm i at their mean values.

```
. mfx, at (median)

Marginal effects after logit

      y  = Pr(y) (predict)

         =  .61688136

------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0018578      .00066    2.82   0.005   .000568  .003147   .780464
      x2 |    .004913       .0037    1.33   0.184   -.00233  .012156   .549624
      x3 |   .0484171      .01963    2.47   0.014   .009946  .086888   .174391
------------------------------------------------------------------------------
```

c. Perform hypothesis testing and make conclusion of the test

. *LR test (Overall test)

```
. logit y

Iteration 0:   log likelihood = -124.34324

Iteration 1:   log likelihood = -124.34324

Logistic regression                               Number of obs     =        215
                                                  LR chi2(0)        =       0.00
                                                  Prob > chi2       =          .
Log likelihood = -124.34324                       Pseudo R2         =     0.0000


------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   1.019544   .1545088     6.60   0.000     .7167121    1.322375
------------------------------------------------------------------------------


. est store logit_res

. lrtest logit logit_res

Likelihood-ratio test                                 LR chi2(3)  =     22.04
(Assumption: logit_res nested in logit)               Prob > chi2 =    0.0001
```

From the LR test above as the p-value is 0.0001 < 0.05, we can reject the null hypothesis that the estimated results are overall insignificant. All the estimated results are overall significant.

Mai Surintraboon
9_6204640178

```
. *Wald test (Overall test)
. test x1 x2 x3

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

           chi2(  3) =   16.43
         Prob > chi2 =    0.0009
```

From the p-value of 0.0009 < 0.05, we can reject the null hypothesis that the estimated results are overall insignificant. All the estimated results are overall significant.

```
. *LR test beta1=beta2=0
. logit y x3
Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -119.39579
Iteration 2:   log likelihood = -119.24145
Iteration 3:   log likelihood = -119.24141
Iteration 4:   log likelihood = -119.24141
Logistic regression                             Number of obs     =        215
                                                LR chi2(1)        =      10.20
                                                Prob > chi2       =     0.0014
Log likelihood = -119.24141                     Pseudo R2         =     0.0410

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x3 |   .2126672   .0754388     2.82   0.005     .0648098    .3605247
       _cons |    .756226   .1716714     4.41   0.000     .4197563    1.092696
------------------------------------------------------------------------------
. est store logit_res1
. lrtest logit logit_res1

Likelihood-ratio test                                 LR chi2(2)  =     11.84
(Assumption: logit_res1 nested in logit)              Prob > chi2 =    0.0027
```

From this LR test whether beta1=beta2=0, as the p-value is 0.0027, which is less than 0.05. We can reject the null hypothesis that beta1=beta2=0. Thus, beta1 and beta are both significant.

d. Why is threshold in computing Counted R2 important? If we change the threshold, can the value of counted R2change? Why? (2 points)

The threshold is important in the determination of the type I and type II errors. Different threshold places different importance on these errors. For example, when using the 0.5 threshold of predicted value, it means that we place equal importance on the type I and II errors.

If we change the value of the threshold, R2 can changed. In placing different importance on type I and II errors, the same model can show out different goodness of fit. It may fit more or fit less.

Question 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)

```
. xtset id t

       panel variable:  id (strongly balanced)

        time variable:  t, 1 to 12

                delta:  1 unit


. 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

------------------------------------------------------------------------------
```

Mai Surintraboon
9_6204640178

```
. 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
```

The null hypothesis of no heteroskedasticity cannot be rejected as the p-value is 0.6629 > 0.05. Therefore, there exist insignificant heteroskedasticity.

If there occurs to be heteroskedasticity in the model, we should use the command:…,vce(robust) when running the regression. The command will be using the robust variance formula, which could solve heteroskedasticity.

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
```

Mai Surintraboon
9_6204640178

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

```
. est store fixed
```

Random-effects 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 random
. hausman random fixed
```

Mai Surintraboon
9_6204640178

	---- Coefficients ----			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	random	fixed	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

In choosing the most appropriated model, we look at 2 tests: the Fixed effects test and the Hausman test.

According to the fixed effects test at the fixed effects model (red color), the p-value = 0.000 < 0.05. It means that the null hypothesis that there are no fixed effects can be rejected. Thus, there exist significant fixed effects in this case. We should not employ the Panel least squares model.

Moreover, according to the Hausman test, the p-value is 0.000<0.05. The null hypothesis (H0: the coefficients of the fixed effects and random effects model are the same) can be rejected. Therefore, the fixed effects model is more appropriated than the random effects model.

All in all, among the 3 models, the fixed effects model is the most appropriated one.

Making interpretation of the estimated model

We will now interpret the fixed effect model as it is best in this case (other models use the same way of interpretation).

When X1 increases by 1 unit, y will increase by approximately 21.63%.

When X2 increases by 1 unit, y will increase by approximately 102.86%.

When X3 increases by 1 unit, y will decrease by approximately 47.79%.

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 the Fixed effects and First difference estimation method is that the First difference can be used only with balanced panel data -- the data in which each cross-sectional unit has a complete set of time.

Fixed effects can be used with both balanced and unbalanced panel data: the data in which each cross-sectional unit does not have a complete set of time.

Mai Surintraboon
9_6204640178

The difference between Fixed effects model and Random effects model is that Fixed effects assume that fixed effects are 100% correlated with the independent variables. Through the estimation method, the Fixed effects will be eliminated entirely to maintain the unbiasedness of the model.

However, the random effects model does not assume that the fixed effects are 100% correlated with the independent variables. It may be mildly correlated. Thus, in the estimation method, it may not get rid of the fixed effects entirely.

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)

Normally, it is preferable to use the overall R-squares for simplicity (because yhat is what we want to estimate). It is comparable with our traditional R-squares.

However, we might want to look at the within R-squares as well when we estimate the fixed effects model. We not usually look at between R-squares. The below is the prove obtain from what Ajarn teaches in class.