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

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

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

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

```
reg y1 y2 x3
```

```
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(2, 47)        =    102.32
       Model |   51317.765         2  25658.8825   Prob > F        =    0.0000
    Residual |   11786.735        47  250.781595   R-squared       =    0.8132
-------------+----------------------------------   Adj R-squared   =    0.8053
       Total |     63104.5        49  1287.84694   Root MSE        =    15.836

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          y2 |    .000951   .0001927     4.94   0.000     .0005634    .0013386
          x3 |   .0073134   .0010989     6.66   0.000     .0051026    .0095241
       _cons |   45.89196   16.95574     2.71   0.009     11.78141     80.0025
------------------------------------------------------------------------------
```

```
. reg y2 y1 x1 x2
```

```
      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     30.40
       Model |  8.1551e+09         3  2.7184e+09   Prob > F        =    0.0000
    Residual |  4.1129e+09        46  89411768.2   R-squared       =    0.6647
-------------+----------------------------------   Adj R-squared   =    0.6429
       Total |  1.2268e+10        49   250368037   Root MSE        =    9455.8

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          y1 |   304.4776   46.94277     6.49   0.000     209.9867    398.9685
          x1 |   242.5322   131.7844     1.84   0.072    -22.73601    507.8004
          x2 |  -.0004598   .0003269    -1.41   0.166    -.0011178    .0001981
       _cons |  -37186.64   9151.027    -4.06   0.000     -55606.7   -18766.58
------------------------------------------------------------------------------
```

From structural form, if we substituted (2) in to (1), we would see that error term from equation (2) is correlated with y1. Y1 is an independent variable in (2). Hence, this will cause simultaneity bias. The same thing would occur if we substituted (1) into (2).

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)

```
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    16.14972    0.8057      89.42   0.0000
y2                    50       3     9460.74    0.6644      24.15   0.0000
------------------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0012152   .0005175     2.35   0.021     .0001875    .0022429
          x3 |   .0063034   .0021463     2.94   0.004     .0020413    .0105655
       _cons |   52.80088   21.34905     2.47   0.015     10.40589    95.19587
-------------+----------------------------------------------------------------
y2           |
          y1 |   294.1639   60.84288     4.83   0.000      173.342    414.9859
          x1 |   259.5341   146.4599     1.77   0.080     -31.3062    550.3744
          x2 |  -.0004723   .0003304    -1.43   0.156    -.0011284    .0001838
       _cons |  -35838.91   10458.16    -3.43   0.001    -56606.74   -15071.08
------------------------------------------------------------------------------
Endogenous variables:  y1 y2
Exogenous variables:   x1 x2 x3
------------------------------------------------------------------------------

. *estimate reduced form model

. reg y1 x1 x2 x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     43.68
       Model |  46708.1821         3   15569.394   Prob > F        =    0.0000
    Residual |  16396.3179        46  356.441694   R-squared       =    0.7402
-------------+----------------------------------   Adj R-squared   =    0.7232
       Total |     63104.5        49  1287.84694   Root MSE        =     18.88

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .5194358   .2532867     2.05   0.046     .0095962    1.029275
          x2 |  -5.53e-07   6.48e-07    -0.85   0.398    -1.86e-06    7.51e-07
          x3 |   .0096401   .0011705     8.24   0.000     .0072841    .0119962
       _cons |   14.25755   19.58637     0.73   0.470    -25.16777    53.68286
------------------------------------------------------------------------------

. predict y1hat
(option xb assumed; fitted values)

. reg y2 x1 x2 x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     17.20
       Model |  6.4858e+09         3  2.1619e+09   Prob > F        =    0.0000
    Residual |  5.7823e+09        46   125701565   R-squared       =    0.5287
-------------+----------------------------------   Adj R-squared   =    0.4979
       Total |  1.2268e+10        49   250368037   Root MSE        =     11212

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

```
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   412.3334    150.414     2.74   0.009     109.5656    715.1012
          x2 |   -.000635   .0003846    -1.65   0.106    -.0014092    .0001391
          x3 |   2.835783   .6950862     4.08   0.000     1.436648    4.234919
       _cons |  -31644.85   11631.35    -2.72   0.009    -55057.54   -8232.164
------------------------------------------------------------------------------

. predict y2hat
(option xb assumed; fitted values)

. *estimate structural form

. *(1)

. reg y1 y2hat x3

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(2, 47)        =     66.60
       Model |  46645.5376         2  23322.7688   Prob > F        =    0.0000
    Residual |  16458.9624        47  350.190689   R-squared       =    0.7392
-------------+----------------------------------   Adj R-squared   =    0.7281
       Total |     63104.5        49  1287.84694   Root MSE        =    18.713

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y2hat |   .0012152   .0005997     2.03   0.048     8.78e-06    .0024216
          x3 |   .0063034    .002487     2.53   0.015     .0013002    .0113066
       _cons |   52.80088   24.73808     2.13   0.038      3.03428    102.5675
------------------------------------------------------------------------------

. *(2)

. reg y2 y1hat x1 x2

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     17.20
       Model |  6.4858e+09         3  2.1619e+09   Prob > F        =    0.0000
    Residual |  5.7823e+09        46   125701569   R-squared       =    0.5287
-------------+----------------------------------   Adj R-squared   =    0.4979
       Total |  1.2268e+10        49   250368037   Root MSE        =     11212

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y1hat |   294.1639   72.10328     4.08   0.000     149.0275    439.3003
          x1 |   259.5341   173.5657     1.50   0.142    -89.83556    608.9038
          x2 |  -.0004723   .0003915    -1.21   0.234    -.0012605    .0003158
       _cons |  -35838.91   12393.69    -2.89   0.006     -60786.1   -10891.71
------------------------------------------------------------------------------
```

Reduced form model are 1. $Y_{1t} = \Pi_{10} + \Pi_{11}x1 + \Pi_{12}x2 + \Pi_{13}x3 + w_{1t}$ and 2. $Y_{2t} = \Pi_{20} + \Pi_{21}x1 + \Pi_{22}x2 + \Pi_{23}x3 + w_{2t}$.

Structural form are

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

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

Endogenous variables are Y_{1t} and Y_{2t}. Exogenous variables are x1 x2 x3.

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    15.65773    0.8057     190.26    0.0000
y2                    50       3    9089.089    0.6633      78.97    0.0000
------------------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          y2 |   .0012152   .0005018     2.42   0.015     .0002317    .0021986
          x3 |   .0063034   .0020809     3.03   0.002     .0022249    .0103819
       _cons |   52.80088   20.69868     2.55   0.011     12.23222    93.36954
-------------+----------------------------------------------------------------
y2           |
          y1 |   289.7893   57.71319     5.02   0.000     176.6735    402.9051
          x1 |   268.9048   139.2509     1.93   0.053    -4.021861    541.8316
          x2 |  -.0003903   .0002722    -1.43   0.152    -.0009239    .0001432
       _cons |  -35811.05   10030.96    -3.57   0.000    -55471.37   -16150.73
------------------------------------------------------------------------------

Endogenous variables:  y1 y2
Exogenous variables:   x1 x2 x3
```

OLS only employ least square method. 2SLS employ IV technique to get rid of correlation between error term and independent variables if there were. 3SLS adds generalize least square method after employing IV technique.

In case of simultaneous equation, OLS will be both biased and inconsistent because exogeneity assumption is violated. To solve the problem, we can employ 2sls and use exogenous variables as instrumental variable to get rid of the correlation between error term and independent variables. However, if there existed correlation of the error term across equations, 3sls will be more asymptotically efficient.

2. From the data set "Final_q2_no.dta":

The study of cost function assumes the model follows CES cost function:

$$c = \lambda\left[\theta R^{-\alpha} + (1-\theta)W^{-\alpha}\right]^{-\frac{\beta}{\alpha}} + \varepsilon \qquad (3)$$

where:

c = Total cost
R = Capital cost
W = Labor cost
λ = Efficiency Parameter
θ = Distribution Parameter
β = Parameter
α = Substitution Parameter
ε = Disturbance Term

Elasticity of Substitution can be computed from: $\sigma = \dfrac{1}{1-\theta}$

The model can then be transformed as:

$$\ln c = \ln\lambda - \frac{\beta}{\alpha}\ln\left[\theta R^{-\alpha} + (1-\theta)W^{-\alpha}\right] + \varepsilon \qquad (4)$$

a. Estimate the model (4) using NLS estimation method using initial values of $\ln\lambda$=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 = {lnlambda}-({beta}/{alpha})*ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha}))), init(lnlambda 1 theta 0.5 beta 0.5 alpha -0.5) nolog
(obs = 252)

      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]
-------------+----------------------------------------------------------------
   /lnlambda |   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 lnlambda taken as constant term in model & ANOVA table
. est store lognls
```

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

. sca list sigma
     sigma =  1.2662762
sca lambda = exp(_b[/lnlambda])

. sca list lambda
    lambda =  7.3624037

. 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
```

Estimate value of lambda is 7.3624037, distribution value(theta) is .2102829, beta is .8545114, alpha is -.931979, and elasticity of substitution is 1.2662762.
From F-test, it suggests that theta, alpha, and beta are not equal to 0.

b. What will happy if we change initial values to $\ln\lambda$=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 = {lnlambda}-({beta}/{alpha})*ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha}))), init(lnlambda 0.5 theta 0.1 beta 0.1 a
> lpha -0.1) nolog
(obs = 252)


      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]
-------------+----------------------------------------------------------------
   /lnlambda |   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 lnlambda taken as constant term in model & ANOVA table
```

If we compared result from (a), we would see that nothing has changed. This is due to log form gives robust estimator regardless of the initial value.

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.

. nl (lnC = {lnlambda}-({beta}/{alpha})*ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha}))), init(lnlambda 0.5 theta 0.1 beta 0.1 a
> lpha -0.1) eps(1e-1) nolog
(obs = 252)


      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]
-------------+----------------------------------------------------------------
   /lnlambda |   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

. drop _est_lognls1

. est store lognls2
*II.

. nl (lnC = {lnlambda}-({beta}/{alpha})*ln({theta}*R^(-{alpha})+(1-{theta})*W^(-
{alpha}))), init(lnlambda 0.5 theta 0.1 beta 0.1 a
> lpha -0.1) eps(1e-15) iter(40) nolog
(obs = 252)


      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]
-------------+----------------------------------------------------------------
   /lnlambda |   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
------------------------------------------------------------------------------
```

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

```
est table lognlsb lognls2 lognls3, star(.1 .05 .01)stat(N rss r2 r2_a)
```

Variable	lognlsb	lognls2	lognls3
lnlambda			
_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

We prefer to use log model because it provides more robust estimation. As we can see from the table, the estimation rarely change when we change iter() or eps().

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

Index function for the decision model can be stated as.

$$I_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \tag{5}$$

The log-likelihood function of this model is as follows:

$$\ln L = \begin{cases} \ln \Phi(z_i) \; if \; y_i = 1 \\ \ln \Phi(-z_i) \; if \; y_i = 0 \end{cases} \tag{6}$$

where: y_i = 1 for yes and 0 for no.

x_{1i} = independent variable 1 of individual i

x_{2i} = independent variable 2 of individual i

Let $\Phi(.)$ = Logistic probability distribution function. $\Phi(z_i) = \dfrac{1}{1+e^{-z_i}}$

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)

```
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 NR

. 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
```

we can see that there is slight difference between models. This is because different computation function used in each model.

b. Perform hypothesis testing whether $\beta_1 = \beta_2 = 0$ using LR-test and Wald test. Make comparison between the two tests. Which test is preferable? Why? (5 points)

```
    est restore NR
(results NR are active now)

. test x1 x2

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

           chi2(  2) =   66.43
         Prob > chi2 =    0.0000

. *LR test

. *get restricted model

. ml maximize
you must issue -ml model- first
r(198);

. 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 NR res

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

P-value from both tests is 0.000. Ho is rejected. Hence, Beta 1 and 2 are not zero. Normally, LR test is preferred because it gives optimal power (using 2 models).

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)

Overall test is a chi-square because the formula for the test gives non-negative value.

F-test cannot be employed because MLE test does not give RSS or R-squared which are the essential elements in performing F-test.

Let $\Phi(.)$ = Cumulation standard normal probability distribution function and

$$z_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \quad (7)$$

Assume that there exists heteroskedasticity in the model as: $\sigma_i^2 = \exp(\gamma x_{3i})^2$, then,

$\Phi(.)$ = Cumulation standard normal probability distribution function $\Phi\ z_i / \exp(\gamma x_{3i})$

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)

```
    *d

.       do      "C:\Users\Jilllin\OneDrive\Desktop\Thammasat\EE426\12\Midterm_q3      -
ml_probit_het.do"

. 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 =  -254.5202
Iteration 3:   log likelihood = -234.73303
Iteration 4:   log likelihood = -230.43623  (not concave)
Iteration 5:   log likelihood = -228.11956
Iteration 6:   log likelihood = -225.58505
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    .7125441
------------------------------------------------------------------------------

. est store hetprob

. do "C:\Users\Jilllin\OneDrive\Desktop\Thammasat\EE426\12\Midterm_q3 - ml_probit.do"

. 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 = -292.02536
rescale:       log likelihood = -275.05597
Iteration 0:   log likelihood = -275.05597
Iteration 1:   log likelihood = -228.81014
Iteration 2:   log likelihood = -228.59561
Iteration 3:   log likelihood = -228.59552
Iteration 4:   log likelihood = -228.59552

                                                Number of obs     =        400
                                                Wald chi2(2)      =      77.26
Log likelihood = -228.59552                     Prob > chi2       =     0.0000
```

```
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .1428731   .0655991     2.18   0.029     .0143013     .271445
          x2 |  -.3256139   .0381417    -8.54   0.000    -.4003703   -.2508575
       _cons |   .3642278   .1170545     3.11   0.002     .1348053    .5936503
------------------------------------------------------------------------------

. est store probit

. lrtest hetprob probit

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

P-value from LR test is less than 0.05. Ho is rejected. Hence, there exists significant heteroskedasticity.

In this case, Wald and LM test cannot be performed since we need 2 models to compare each other.

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)

In MLE, distribution of the estimated parameters are not always t-distributed depending on distribution of the regression model. Thus, individual test can be performed by Z-test.

We do not have R-square in MLE because the process of getting estimated result is comes from maximizing log-likelihood function, not fitting the line with the data we obtained. If we plotted the model, the axis would be log-likelihood value and estimated parameter.

Hence, to compare across models, log-likelihood value is more appropriate. The model with higher log-likelihood value would suggest that it is better.

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

$$r_{t+\Delta t} - r_t = (\alpha + \beta r_t)\Delta t + \varepsilon_{t+\Delta t} \quad (8)$$

where: $E[\varepsilon_{t+\Delta t}] = 0$ and $E[\varepsilon_{t+\Delta t}^2] = \sigma^2 r_t^{2\gamma} \Delta t$

Then, the model can be transformed to be discrete time model by setting $\Delta t = 1$.
The discrete time model can be stated as:

$$r_{t+1} - r_t = \Delta r_t = \alpha + \beta r_t + \varepsilon_{t+1} \quad (9)$$

where: $E[\varepsilon_{t+1}] = 0$ and $E[\varepsilon_{t+1}^2] = \sigma^2 r_t^{2\gamma}$

The model consists of four parameters, including $\alpha, \beta, \sigma^2, \gamma$. Four moment condition equations of the model can be stated as:

(1) Zero mean condition: $E(\varepsilon_{t+1}) = 0$

(2) Orthogonality condition: $E(\varepsilon_{t+1} r_t) = 0$

(3) Variance condition: $E(\varepsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}) = 0$

(4) Zero covariance condition: $E\left(\left(\varepsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}\right) r_t\right) = 0$

The above model can be claimed as unrestricted model for other two interest rate structure models which can be indicated as follows:

Model	α	β	σ^2	γ
(1) Unrestricted				
(2) Merton		0		0
(3) Vasicek				0

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).

```
use "C:\Users\Jilllin\OneDrive\Desktop\Thammasat\EE426\12\Midterm_q4-1_12.dta", clear

. 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) =  .00001191
Iteration 1:   GMM criterion Q(b) =  8.722e-06  (backed up)
Iteration 2:   GMM criterion Q(b) =  6.129e-06  (not concave)
Iteration 3:   GMM criterion Q(b) =  5.573e-06  (backed up)
Iteration 4:   GMM criterion Q(b) =  5.489e-06  (backed up)

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

note: model is exactly identified

GMM estimation

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

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

. est store UR

. *Merton

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

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

```
Iteration 2:   GMM criterion Q(b) =  4.144e-08

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

GMM estimation

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

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

. est store Merton

. *Vasicek

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

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

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

GMM estimation

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

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

```
-------------+----------------------------------------------------------------
      /alpha |  -.0027101   .0009776    -2.77   0.006    -.0046262    -.000794
       /beta |   .0005363   .0001997     2.69   0.007     .0001449    .0009277
     /sigma2 |   .0005959   .0003005     1.98   0.047     6.90e-06     .001185
------------------------------------------------------------------------------
Instruments for equation 1: _cons
Instruments for equation 2: _cons
Instruments for equation 3: _cons
Instruments for equation 4: _cons

. est store Vasicek

est table UR Merton Vasicek, star(.1 .05 .01) stat(N J)

----------------------------------------------------------
    Variable |      UR            Merton           Vasicek
-------------+--------------------------------------------
alpha        |
       _cons | -.00239168**     -.0008227      -.00271012***
-------------+--------------------------------------------
beta         |
       _cons |  .00043212                       .00053631***
-------------+--------------------------------------------
sigma2       |
       _cons |  .00051327       .00043542       .00059595**
-------------+--------------------------------------------
gamma        |
       _cons |   .0942137
-------------+--------------------------------------------
Statistics   |
           N |       1325            1325            1325
           J |  2.454e-21       7.3506315       .25544087
----------------------------------------------------------
                          legend: * p<.1; ** p<.05; *** p<.01
. est restore Merton
(results Merton are active now)

. estat overid

  Test of overidentifying restriction:

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

From overid test of Merton, Ho is rejected, suggesting that the conditions are not true under Merton model.

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

```
    *b

. est restore UR
(results UR are active now)

. *test Merton

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

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

           chi2(  2) =    7.89
         Prob > chi2 =    0.0194

. *test Vasicek

. test (_b[/gamma]=0)

 ( 1)  [gamma]_cons = 0

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

From Wald test, it suggests that Vasicek it the most appropriate model. The criteria used is employing Wald test. We can see that we cannot reject the null hypothesis of gamma = 0 which conform with Vasicek model. On the other hand, we can reject the null hypothesis of Merton model, suggesting that the model is not true.

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

The model: $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i$ (10)

where: $u_i \sim N(0, \sigma^2)$, $E(x_{3i} u_i) = 0$ and $E(x_{4i} u_i) = 0$ but $E(x_{1i} u_i) \neq 0$ and $E(x_{2i} u_i) \neq 0$

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

The model will be both biased and inconsistent because we violate exogeneity assumption.

If $E(z_{1i} u_i) = 0$, $E(z_{2i} u_i) = 0$, $E(z_{3i} u_i) = 0$, and z_{1i}, z_{2i}, and z_{3i} are highly correlated with x_{1i} and x_{2i}

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

```
use     "C:\Users\Jilllin\OneDrive\Desktop\Thammasat\EE426\12\Midterm_q4-2_12.dta",
clear

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

Instrumental variables (GMM) regression          Number of obs   =        500
                                                 Wald chi2(4)    =     161.20
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.6026
GMM weight matrix: Robust                        Root MSE        =     20.049

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

```
------------+----------------------------------------------------------------
          x1 |    3.00695   .9756557     3.08   0.002       1.0947      4.9192
          x2 |   2.154062   .5140463     4.19   0.000      1.14655    3.161574
          x3 |   1.059671   .3118073     3.40   0.001     .4485398    1.670802
          x4 |    1.32458   .1765816     7.50   0.000     .9784861    1.670673
       _cons |   2.062215   7.475435     0.28   0.783    -12.58937     16.7138
------------------------------------------------------------------------------
Instrumented:  x1 x2
Instruments:   x3 x4 z1 z2 z3

. estat overid

  Test of overidentifying restriction:

  Hansen's J chi2(1) = .003095 (p = 0.9556)

. estat endogenous x1 x2

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

  GMM C statistic chi2(2) =   173.88  (p = 0.0000)

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

Instrumental variables (2SLS) regression          Number of obs   =        500
                                                  Wald chi2(4)    =     150.72
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.6026
                                                  Root MSE        =     20.048

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   3.006552    .954753     3.15   0.002      1.13527    4.877833
          x2 |   2.154759   .5162653     4.17   0.000     1.142898    3.166621
          x3 |   1.060662   .3116292     3.40   0.001     .4498799    1.671444
          x4 |   1.324314   .1795983     7.37   0.000     .9723082     1.67632
       _cons |   2.047395   8.119868     0.25   0.801    -13.86725    17.96204
------------------------------------------------------------------------------
Instrumented:  x1 x2
Instruments:   x3 x4 z1 z2 z3
```

From overid test, we cannot reject Ho which is indicative of overidentification. Furthermore, endogenous test also suggests that there exists endogeneity bias from x1 and x2. Hence, these tests confirm that using GMM is appropriate.

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

Difference between GMM and 2SLS is the approach. GMM uses moment conditions while 2SLS uses IV technique. In this case, there are 3 instrumental variables for 2 independent variables, causing GMM to use more condition than 2SLS.

2SLS will run each Xs with the IVs, so there are 2 equations.

GMM use moment condition for each IVs, so there are 3 equations.

On this ground, GMM will be more efficient because that it utilizes more data.

5. From the data set "Midterm_q5_no.dta":
In the study of decision to pay dividend of listed company in the stock market, the study states the following probability function:

$$I_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} \tag{11}$$

The log-likelihood function of this model is as follows:

$$\ln L = \begin{cases} \ln \Lambda(I_i) \text{ if } y_i = 1 \\ \ln \Lambda(-I_i) \text{ if } y_i = 0 \end{cases} \tag{12}$$

where: y_i = 1 for dividend paying firm and 0 otherwise.
x_{1i} = Liquidity ratio of firm i
x_{2i} = log of firm size firm i
x_{3i} = Profitability index of firm i

(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)

```
use "C:\Users\Jilllin\OneDrive\Desktop\Thammasat\EE426\12\Midterm_q5_12.dta", clear

. probit y x1 x2 x3

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -113.56856
Iteration 2:   log likelihood = -113.39876
Iteration 3:   log likelihood = -113.39854
Iteration 4:   log likelihood = -113.39854

Probit regression                                 Number of obs     =        215
                                                  LR chi2(3)        =      21.89
                                                  Prob > chi2       =     0.0001
Log likelihood = -113.39854                       Pseudo R2         =     0.0880

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0045224   .0014463     3.13   0.002     .0016878     .007357
          x2 |   .0103217   .0082387     1.25   0.210    -.0058259    .0264692
          x3 |   .1123308   .0428477     2.62   0.009     .0283509    .1963107
       _cons |   .2708915   .1195254     2.27   0.023      .036626    .5051569
------------------------------------------------------------------------------

. est store probit

. logit y x1 x2 x3

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -114.01642
```

```
Iteration 2:   log likelihood = -113.45223
Iteration 3:   log likelihood = -113.45017
Iteration 4:   log likelihood = -113.45017

Logistic regression                             Number of obs     =        215
                                                LR chi2(3)        =      21.79
                                                Prob > chi2       =     0.0001
Log likelihood = -113.45017                     Pseudo R2         =     0.0876

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0077938   .0026047     2.99   0.003     .0026886     .012899
          x2 |   .0197148   .0155308     1.27   0.204    -.0107251    .0501546
          x3 |    .201728   .0802515     2.51   0.012     .0444379    .3590181
       _cons |   .4283187   .1939944     2.21   0.027     .0480967    .8085406
------------------------------------------------------------------------------

. fitstat
command fitstat is unrecognized
r(199);

. findit fitstat

. fitstat

Measures of Fit for logit of y

Log-Lik Intercept Only:     -124.343     Log-Lik Full Model:         -113.450
D(211):                      226.900     LR(3):                        21.786
                                         Prob > LR:                     0.000
McFadden's R2:                 0.088     McFadden's Adj R2:             0.055
Maximum Likelihood R2:         0.096     Cragg & Uhler's R2:            0.141
McKelvey and Zavoina's R2:     0.191     Efron's R2:                    0.089
Variance of y*:                4.068     Variance of error:             3.290
Count R2:                      0.735     Adj Count R2:                  0.000
AIC:                           1.093     AIC*n:                       234.900
BIC:                        -906.304     BIC':                         -5.674

. est store logit
```

We cannot compare across models because both models assume different probability distribution.

From fitstat and estimated result, overall test is significant. For individual test, it suggests that X2 is not significant in this model. Pseudo-R^2 is 0.0876 and Counted R^2 is 0.735.

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

```
est restore logit
(results logit are active now)

. mfx

Marginal effects after logit
      y  = Pr(y) (predict)
```

```
         =  .76841074
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.      z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |    .001387      .00044    3.16   0.002   .000528  .002246   60.8652
      x2 |   .0035084      .00271    1.29   0.195  -.001802  .008819  -1.64234
      x3 |   .0358986      .01356    2.65   0.008   .009317   .06248   1.63115
------------------------------------------------------------------------------

. mfx, at(median)

Marginal effects after logit
      y  = Pr(y) (predict)
         =  .61730017
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.      z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0018412      .00065    2.81   0.005   .000559  .003124   .780464
      x2 |   .0046574      .00368    1.26   0.206  -.002564  .011879   .468091
      x3 |   .0476564      .01952    2.44   0.015   .009406  .085907   .170897
------------------------------------------------------------------------------

.
```

For marginal effect at mean of X_{1t}, if liquidity ratio of firm i increase by 1, it would increase the probability of the firm to pay dividend by .001387.

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

```
*c

. *Wald test

. test x1=x2=x3=0

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

           chi2(  3) =   16.36
         Prob > chi2 =    0.0010

. *LR test

. *get restricted model

. 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 res

. lrtest logit res

Likelihood-ratio test                                 LR chi2(3)  =     21.79
(Assumption: res nested in logit)                     Prob > chi2 =    0.0001

. lrtest help
estimation result help not found
r(111);

. help lrtest

. logit y x1 x3

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -114.65962
Iteration 2:   log likelihood = -114.26202
Iteration 3:   log likelihood = -114.25999
Iteration 4:   log likelihood = -114.25999

Logistic regression                             Number of obs     =        215
                                                LR chi2(2)        =      20.17
                                                Prob > chi2       =     0.0000
Log likelihood = -114.25999                     Pseudo R2         =     0.0811

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0071152   .0024596     2.89   0.004     .0022944     .011936
          x3 |   .1758473   .0770325     2.28   0.022     .0248664    .3268282
       _cons |   .4536892   .1926964     2.35   0.019     .0760113    .8313671
------------------------------------------------------------------------------

. est store res1

. logit y x2 x3

Iteration 0:   log likelihood = -124.34324
Iteration 1:   log likelihood = -119.19101
Iteration 2:   log likelihood = -119.00812
Iteration 3:   log likelihood = -119.00806
Iteration 4:   log likelihood = -119.00806

Logistic regression                             Number of obs     =        215
                                                LR chi2(2)        =      10.67
                                                Prob > chi2       =     0.0048
Log likelihood = -119.00806                     Pseudo R2         =     0.0429

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

```
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |   .0094376   .0128611     0.73   0.463    -.0157697    .0346448
          x3 |   .2235154   .0774376     2.89   0.004     .0717404    .3752903
       _cons |   .7562891   .1720119     4.40   0.000     .4191519    1.093426
------------------------------------------------------------------------------

. est store res2

. lrtest logit res1

Likelihood-ratio test                                 LR chi2(1)  =      1.62
(Assumption: res1 nested in logit)                    Prob > chi2 =    0.2031

. lrtest logit res2

Likelihood-ratio test                                 LR chi2(1)  =     11.12
(Assumption: res2 nested in logit)                    Prob > chi2 =    0.0009
```

Both Wald and LR test suggest that β_1, β_2, and β_3 are not equal to 0.

To test if $\beta_1 = \beta_2$, 2 restricted models are generated, LR test of both models with the unrestricted model are not equal. Hence, it suggests that beta1 and beta2 are not equal.

(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)

The threshold is important because it would determine the probability of getting false positive or false negative. Hence, it is critically important when we do certain evaluation. For example, if we are evaluating if a patient is sick or not, we would want the threshold to be low because getting false positive is better than getting false negative.

Counted R-squared is the ratio of correct estimation of the model over the total observation. Changing threshold would more or less change the estimated result. Hence, the value of counted R-squared can be changed according to the threshold.

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

Panel Data Model

$$y_{it} = \alpha + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + u_{it} \quad (13)$$

where: y_{it} is Dependent variable.
x_{kit} is Independent variable k.
u_{it} is Stochastic disturbance term.

Fixed Effects Model

$$y_{it} = \alpha_i + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + u_{it} \quad (14)$$

where: α_i is Cross-sectional fixed effects

Random Effects Model

$$y_{it} = \alpha + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + u_{it} \quad (15)$$

and $u_{it} = v_i + \varepsilon_{it}$

where: v_i = Cross-section random effects
ε_{it} = residual terms

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)

```
use "C:\Users\Jilllin\OneDrive\Desktop\Thammasat\EE426\12\Midterm_q6_12.dta", clear

. xtset id t
       panel variable:  id (strongly balanced)
        time variable:  t, 1 to 12
                delta:  1 unit

. *a

. 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 hetero

. 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 pls

. local df = e(N_g) - 1

. lrtest hetero, df(`df')

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

From LR test, we cannot reject Ho. It is indicative of not having heteroskedasticity problem. In this case, POLS would be a viable option since there is no correlation between error terms.

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)

```
*b

. xtgls y x1 x2 x3
```

```
Cross-sectional time-series FGLS regression

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

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

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

. *fixed effect model

. xtgls y x1 x2 x3, fe
option fe not allowed
r(198);

. xtgls y x1 x2 x3,fe
option fe not allowed
r(198);

. 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

. 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 fixed re

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

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       re         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

.
```

From FE model, fixed effect test suggests that there exists a significant fixed effect (all α_i are not equal). From Hausman test, the null hypothesis is rejected because p-value is less than 0.05. Hence, fixed effect model is more appropriate.

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)

Fixed effect and first difference estimation share the same idea of getting rid of the unobserved fixed effect (α_i). The difference is the way the get rid of it. For fixed effect, it gets rid of α_i by subtracting time average equation from the original model. Since α_i is not time variated, average of α_i equals to α_i. Hence, α_i is eliminated.

Likewise, first difference model utilizes the same property of α_i. It subtracts the original model with the value of that model in period t-1. According to non-time-variated property of α_i. α_t and α_{t-1} are equal, so α_i is eliminated.

Random effect estimation assumes that $cov(\alpha_i, X_i) = 0$. Subsequently, it employs fGLS technique by using the weight lambda because having α_i in the error term still causes correlation among error terms. Value of lambda is obtained from variance of u and α_i.

However, if variance of α_i is approaching infinity (really large). Random effect estimation would be the same as fixed effect estimation.

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)

```
. 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

From fixed effect test, within R^2 explain the variation of Y in each ID for 95.02 percent. Between R^2 explain the variation of Y across ID for 98.48 percent. Overall R^2 is the weighted average of the two, implying overall ability of the model to explain variation of Y for 88.46 percent.