(a) Estimate the model using multinomial logit model using y_i=0 is the base outcome. Make interpretation of your estimated results (sign & meaning (in term of rrr), overall test, goodness of fit, and individual test). Perform IIA test whether it is appropriated to apply multinomial logit in this case? Give explanation why IIA is important for multinomial logit. Determine whether multinomial logit is appropriated for this case? Why? (5 points).

1.

```
. mlogit y x1 x2 x3 x4, base(0) nolog

Multinomial logistic regression                 Number of obs     =        170

                                                LR chi2(8)        =      94.19

                                                Prob > chi2       =     0.0000

Log likelihood = -100.73601                     Pseudo R2         =     0.3186

------------------------------------------------------------------------------

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

-------------+----------------------------------------------------------------

0            |  (base outcome)

-------------+----------------------------------------------------------------

1            |

          x1 |  -.8716535   .6283126    -1.39   0.165    -2.103124    .3598166

          x2 |  -.5356032   .7697797    -0.70   0.487    -2.044344    .9731372

          x3 |   -.517283   .6320378    -0.82   0.413    -1.756054    .7214882

          x4 |   .5657851   .2625729     2.15   0.031     .0511516    1.080419

       _cons |  -6.672078   3.719218    -1.79   0.073    -13.96161    .6174562

-------------+----------------------------------------------------------------

2            |

          x1 |  -.7377626   .7162596    -1.03   0.303    -2.141606    .6660804

          x2 |  -1.557584   .8090783    -1.93   0.054    -3.143348    .0281805

          x3 |   .6492327   .6826875     0.95   0.342    -.6888103    1.987276

          x4 |   1.886577   .3228789     5.84   0.000     1.253746    2.519408

       _cons |  -27.46135    4.83019    -5.69   0.000    -36.92835   -17.99435

------------------------------------------------------------------------------

. est store m

. mlogit y x1 x2 x3 x4 if y!=2, base(0) nolog

Multinomial logistic regression                 Number of obs     =         60

                                                LR chi2(4)        =       5.96

                                                Prob > chi2       =     0.2023

Log likelihood = -34.480553                     Pseudo R2         =     0.0795

------------------------------------------------------------------------------

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

-------------+----------------------------------------------------------------
```

```
0            |  (base outcome)
-------------+----------------------------------------------------------------
1            |
          x1 |  -.7497427   .6037771    -1.24   0.214    -1.933124    .4336386
          x2 |  -.2864689   .7421089    -0.39   0.699    -1.740976    1.168038
          x3 |  -.3366018    .640531    -0.53   0.599    -1.592019    .9188158
          x4 |   .4928915    .257627     1.91   0.056    -.0120481    .9978311
       _cons |  -5.928091   3.797311    -1.56   0.118    -13.37068    1.514502
------------------------------------------------------------------------------
. est store mno2
. hausman m mno2, alleqs constant
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       m           mno2        Difference          S.E.
-------------+----------------------------------------------------------------
          x1 |   -.8716535    -.7497427       -.1219107        .1738677
          x2 |   -.5356032    -.2864689       -.2491343        .2045363
          x3 |    -.517283    -.3366018       -.1806812               .
          x4 |    .5657851     .4928915        .0728936        .0507236
       _cons |   -6.672078    -5.928091       -.7439869               .
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from mlogit
           B = inconsistent under Ha, efficient under Ho; obtained from mlogit
    Test:  Ho:  difference in coefficients not systematic
                  chi2(5) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        1.05
                Prob>chi2 =      0.9584
                (V_b-V_B is not positive definite)
. mlogit y x1 x2 x3 x4, rrr base(0) nolog
Multinomial logistic regression                 Number of obs     =        170
                                                LR chi2(8)        =      94.19
                                                Prob > chi2       =     0.0000
Log likelihood = -100.73601                     Pseudo R2         =     0.3186
------------------------------------------------------------------------------
           y |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |  (base outcome)
```

```
-------------+----------------------------------------------------------------
1            |
          x1 |   .4182594   .2627976    -1.39   0.165     .1220745    1.433067
          x2 |   .5853161   .4505644    -0.70   0.487     .1294651    2.646233
          x3 |    .596138   .3767818    -0.82   0.413      .172725    2.057493
          x4 |    1.76083   .4623462     2.15   0.031     1.052482    2.945912
       _cons |   .0012658   .0047077    -1.79   0.073     8.64e-07    1.854205
-------------+----------------------------------------------------------------
2            |
          x1 |   .4781826   .3425029    -1.03   0.303     .1174661    1.946592
          x2 |   .2106444   .1704278    -1.93   0.054     .0431381    1.028581
          x3 |   1.914072   1.306713     0.95   0.342     .5021731    7.295631
          x4 |   6.596747   2.129951     5.84   0.000     3.503441    12.42124
       _cons |   1.18e-12   5.72e-12    -5.69   0.000     9.17e-17    1.53e-08
------------------------------------------------------------------------------
```

Ans. As we use y0 as the base case, the overall test is significant. Pseudo R2 is 0.3186. RRR of x1 x2 x3 in y1 is lower than 1, thus, they have negative impact on y. for x4, it has positive impact. In case of y2, RRR is lower than 1 only for x1 and x2, so only x3 and x4 have positive impact on y. However, the individual tests seem to be mostly insignificant. For x4 in both cases, it is significant under 95 percent confidence level.

After testing IIA, it turns out that we accept the null hypothesis. It means that IIA is valid, then we can use multinomial logit. If IIA is not satisfied and we use multinomial logit, the model will be biased as it is the most important assumption of mnlogit where decisions are independent of other alternatives.

(b) Estimate the model using order probit model of yi. Make interpretation of your estimated results (sign & meaning, overall test, goodness of fit, and individual test). Perform the test to determine whether order probit is appropriated in this case? Give explanation why? (5 points).

```
. oprobit y x1 x2 x3 x4, nolog

Ordered probit regression                        Number of obs     =        170
                                                 LR chi2(4)        =      84.22
                                                 Prob > chi2       =     0.0000
Log likelihood = -105.72267                      Pseudo R2         =     0.2848
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.1825779   .2396318    -0.76   0.446    -.6522476    .2870919
          x2 |   -.489389   .2441328    -2.00   0.045    -.9678804   -.0108975
          x3 |   .4860257   .2228889     2.18   0.029     .0491715      .92288
          x4 |   .7169483   .0945822     7.58   0.000     .5315706     .902326
```

```
-------------+----------------------------------------------------------------
       /cut1 |   9.935009   1.433907                      7.124604    12.74542
       /cut2 |   11.20764   1.492805                      8.281792    14.13348
```

```
. tab y

          y |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         19       11.18       11.18
          1 |         41       24.12       35.29
          2 |        110       64.71      100.00
------------+-----------------------------------
      Total |        170      100.00

. g y12=y>0

. g y2=y>1

. probit y12 x1 x2, nolog

Probit regression                                 Number of obs     =        170
                                                  LR chi2(2)        =       0.37
                                                  Prob > chi2       =     0.8294
Log likelihood = -59.345062                       Pseudo R2         =     0.0031

------------------------------------------------------------------------------
         y12 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.0227776   .2875784    -0.08   0.937    -.5864209    .5408658
          x2 |  -.1670744    .281841    -0.59   0.553    -.7194726    .3853239
       _cons |   1.344802   .2666147     5.04   0.000     .8222469    1.867357
------------------------------------------------------------------------------

. est store p12

. probit y2 x1 x2, nolog

Probit regression                                 Number of obs     =        170
                                                  LR chi2(2)        =       7.58
                                                  Prob > chi2       =     0.0226
Log likelihood = -106.58421                       Pseudo R2         =     0.0343

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |    .511485   .2233564     2.29   0.022     .0737145    .9492554
          x2 |  -.3915635   .2199933    -1.78   0.075    -.8227425    .0396155
       _cons |   .4126665   .1989986     2.07   0.038     .0226365    .8026964
```

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

. est store p2

. suest p12 p2

Simultaneous results for p12, p2

                                                 Number of obs     =        170

------------------------------------------------------------------------------

             |               Robust

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

-------------+----------------------------------------------------------------

p12_y12      |

          x1 |  -.0227776   .3181703    -0.07   0.943      -.64638    .6008249

          x2 |  -.1670744   .2860756    -0.58   0.559    -.7277722    .3936235

       _cons |   1.344802   .2586097     5.20   0.000     .8379364    1.851668

-------------+----------------------------------------------------------------

p2_y2        |

          x1 |    .511485   .2388239     2.14   0.032     .0433987    .9795712

          x2 |  -.3915635   .2247024    -1.74   0.081    -.8319722    .0488452

       _cons |   .4126665   .2000799     2.06   0.039      .020517    .8048159

------------------------------------------------------------------------------

. test [p12_y12]x1=[p2_y2]x1

 ( 1)  [p12_y12]x1 - [p2_y2]x1 = 0


           chi2(  1) =    3.30

         Prob > chi2 =    0.0693


. test [p12_y12]x2=[p2_y2]x2

 ( 1)  [p12_y12]x2 - [p2_y2]x2 = 0


           chi2(  1) =    0.78

         Prob > chi2 =    0.3767
```

Ans. for the order probit model, the overall test is significant. Pseudo R2 is 0.2848, the higher the better. The individual test for x1 is insignificant, other Xs are significant under 95 percent confidence level. The cuts don’t include zero in the confidence interval, they are significant; this can imply that there are differences for making decisions and the model may be ordered. From the test above, we run p12 and p2 simultaneously and test whether there are differences of the coefficients. Here we accept null hypothesis (pvalue greater than 0.5) that they are equal, thus, we have one set of beta so ordered probit should be used.

(c) Estimate models for Y1i, Y2i, and Y3i assuming that the probability functions follow separate normal distribution function. Should we use these three separate probit models? Why? (4 points)

```
. probit y1 x*, nolog

Probit regression                               Number of obs     =        550
                                                LR chi2(4)        =      33.87
                                                Prob > chi2       =     0.0000
Log likelihood =  -361.4398                     Pseudo R2         =     0.0448

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.473191   .4060376     3.63   0.000     .6773715     2.26901
          x2 |  -.9429238   .3812976    -2.47   0.013    -1.690253   -.1955943
          x3 |    .671879   .1790825     3.75   0.000     .3208837    1.022874
          x4 |   .2561454   .1857917     1.38   0.168    -.1079996    .6202903
       _cons |   -.807743   .3058901    -2.64   0.008    -1.407277   -.2082095
------------------------------------------------------------------------------

. probit y2 x*, nolog

Probit regression                               Number of obs     =        550
                                                LR chi2(4)        =      15.11
                                                Prob > chi2       =     0.0045
Log likelihood = -360.03405                     Pseudo R2         =     0.0205

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.8482165   .3797737    -2.23   0.026    -1.592559   -.1038737
          x2 |  -.4076769   .3663393    -1.11   0.266    -1.125689     .310335
          x3 |  -.2409677   .1767855    -1.36   0.173    -.5874609    .1055255
          x4 |   .6114268    .190396     3.21   0.001     .2382575    .9845961
       _cons |   .3242658   .3026141     1.07   0.284     -.268847    .9173786
------------------------------------------------------------------------------

. probit y3 x*, nolog

Probit regression                               Number of obs     =        550
                                                LR chi2(4)        =      12.74
                                                Prob > chi2       =     0.0126
Log likelihood = -330.44911                     Pseudo R2         =     0.0189

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
          x1 |  -.5264988   .3908551    -1.35   0.178    -1.292561    .2395631
          x2 |   .8452187   .3800049     2.22   0.026     .1004228    1.590014
          x3 |  -.3860994   .1982681    -1.95   0.051    -.7746977    .0024989
          x4 |  -.3356344   .1826289    -1.84   0.066    -.6935805    .0223118
       _cons |   -.498158   .3167096    -1.57   0.116    -1.118897    .1225813
```

```
Ans. the models are significant in overall, however we need to look at the
correlation among disturbances too. We can run them separately if the disturbances
are not correlated to each other.
```

(d) Estimate models for Y1i, Y2i, and Y3i assuming that the probability functions follow multivariate normal probability distribution function (MV Probit models). Determine whether MVProbit is appropriated. Why? How does the MV Probit models differ from three separate probit models? (6 points)

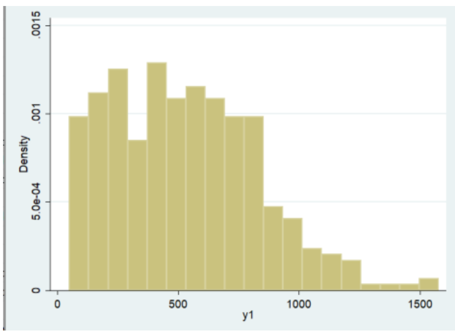
```
. mvprobit (y1 x*) (y2 x*) (y3 x*), nolog

Multivariate probit (MSL, # draws = 5)            Number of obs   =        550
                                                  Wald chi2(12)   =      45.91
Log likelihood = -931.74666                       Prob > chi2     =     0.0000
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y1           |
          x1 |   1.312046   .3955406     3.32   0.001     .5368009    2.087292
          x2 |  -.8140709   .3752078    -2.17   0.030    -1.549465   -.0786772
          x3 |   .6494881   .1798034     3.61   0.000     .2970799    1.001896
          x4 |   .2353198   .1847334     1.27   0.203    -.1267509    .5973906
       _cons |   -.733309   .3074531    -2.39   0.017    -1.335906   -.1307119
-------------+----------------------------------------------------------------
y2           |
          x1 |  -.9163016   .3766512    -2.43   0.015    -1.654524   -.1780789
          x2 |  -.3409062   .3641912    -0.94   0.349    -1.054708    .3728954
          x3 |  -.2596698   .1773654    -1.46   0.143    -.6072996    .0879601
          x4 |   .5935985   .1895573     3.13   0.002      .222073     .965124
       _cons |   .3632508   .3014063     1.21   0.228    -.2274946    .9539963
-------------+----------------------------------------------------------------
y3           |
          x1 |   -.485328   .3837317    -1.26   0.206    -1.237428    .2667723
          x2 |   .6613654   .3751111     1.76   0.078    -.0738389     1.39657
```

```
           x3 |  -.2797045   .1860602    -1.50   0.133    -.6443757    .0849667
           x4 |  -.2873432   .1862661    -1.54   0.123    -.6524182    .0777317
        _cons |  -.4278602   .3180921    -1.35   0.179    -1.051309    .1955888
-------------+----------------------------------------------------------------
    /atrho21 |  -.5589267     .07351    -7.60   0.000    -.7030036   -.4148498
-------------+----------------------------------------------------------------
    /atrho31 |  -.4436816   .0752587    -5.90   0.000     -.591186   -.2961772
-------------+----------------------------------------------------------------
    /atrho32 |  -.3496377   .0713966    -4.90   0.000    -.4895725    -.209703
-------------+----------------------------------------------------------------
       rho21 |  -.5071806   .0546008    -9.29   0.000    -.6062708   -.3925828
-------------+----------------------------------------------------------------
       rho31 |  -.4166914   .0621914    -6.70   0.000     -.530748   -.2878103
-------------+----------------------------------------------------------------
       rho32 |  -.3360542   .0633336    -5.31   0.000     -.453877   -.2066822
------------------------------------------------------------------------------
Likelihood ratio test of  rho21 = rho31 = rho32 = 0:
             chi2(3) =  240.353   Prob > chi2 = 0.0000
```

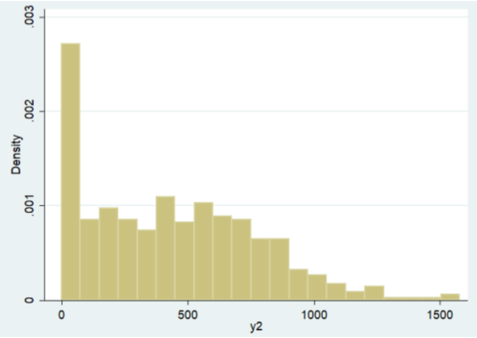
ans. the separate models seem to be significant in overall. However, from test of appropriateness of mvprobit model, it is better to use mvprobit because we reject the null hypothesis that rhos are zero. When rhos are not zero, it means that the disturbances are related and we cannot run them separately. When we run them separately, we assume no correlation among the disturbance terms from each models.

2.(a) Plot histogram of y1i, y2i, y3i, compute descriptive statistics of these three variables. Determine limitations of these three dependent variables. (5 points)

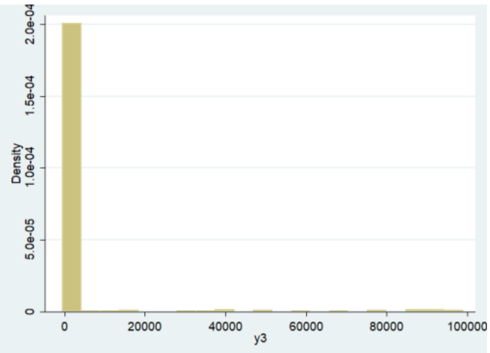
```
. sum y1 y2 y3
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          y1 |        367    528.8534     302.314   50.21759    1578.51
          y2 |        450    432.3137    340.3366          0    1578.51
          y3 |        450    3344.719    14544.57  -466.0042   98951.63
. histogram y1
```



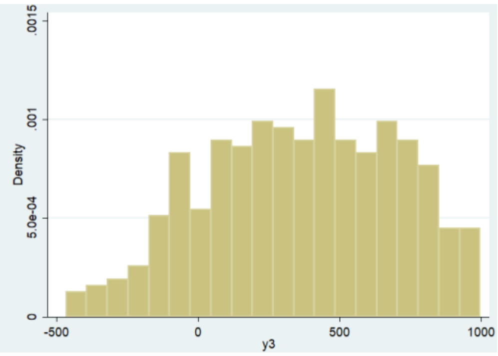
```
. histogram y2
```




```
. histogram y3
```




```
. histogram y3 if y3<2000
```



Ans. from descriptive stats of three variables and histograms, it may be concluded that y1 has truncated problem at its minimum, y2 has censored problem, y3 has outliers problem.

(b) Estimate the model (3) for y1i using OLS and truncated regression model. Determine the most appropriated model in this case. Why? What is the major problem in this case? What will happen if we ignore the problem? (5 points)

```
. reg y1 x

      Source |       SS           df       MS      Number of obs   =       367
-------------+----------------------------------   F(1, 365)       =    139.84
       Model |  9265666.01         1  9265666.01   Prob > F        =    0.0000
    Residual |  24184454.5       365  66258.7795   R-squared       =    0.2770
-------------+----------------------------------   Adj R-squared   =    0.2750
       Total |  33450120.5       366  91393.7719   Root MSE        =    257.41
```

```
------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   166.2833   14.06151    11.83   0.000     138.6316    193.9351
       _cons |  -10.17282   47.52115    -0.21   0.831    -103.6224    83.27678
------------------------------------------------------------------------------
. est store y1
. sum y1
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          y1 |        367    528.8534     302.314   50.21759    1578.51
. scalar miny1=round(r(min))
. truncreg y1 x, ll(miny1) nolog
(note: 0 obs. truncated)
Truncated regression
Limit:   lower =         50                     Number of obs     =        367
         upper =       +inf                     Wald chi2(1)      =     106.88
Log likelihood = -2523.2473                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   230.0903   22.25604    10.34   0.000     186.4693    273.7114
       _cons |  -293.0718   85.01096    -3.45   0.001    -459.6902   -126.4534
-------------+----------------------------------------------------------------
      /sigma |    304.355   16.61638    18.32   0.000     271.7875    336.9225
------------------------------------------------------------------------------
. predict truncated, e(50,.)
. est store y1t
. lrtest y1 y1t, force
Likelihood-ratio test                                 LR chi2(1)  =     67.19
(Assumption: y1 nested in y1t)                        Prob > chi2 =    0.0000
```

Ans. from the test of appropriateness, truncated regression model should be used as we reject the null hypothesis. When we cannot observe the data but it does exits, we are facing truncated problem. If we ignore the truncated problem and run with only available data using ols linear regression model, it will be biased.

(c) Estimate the model (3) for y2i using OLS and Tobit model. Determine the most appropriated model in this case. Why? What is the major problem in this case? What will happen if we ignore the problem? (5 points)

```
. reg y2 x
```

```
      Source |       SS           df       MS      Number of obs   =       450
-------------+----------------------------------   F(1, 448)       =    253.69
       Model |  18802566.3         1  18802566.3   Prob > F        =    0.0000
    Residual |    33204659       448  74117.5425   R-squared       =    0.3615
-------------+----------------------------------   Adj R-squared   =    0.3601
       Total |  52007225.3       449   115829.01   Root MSE        =    272.25
------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   195.9104   12.30012    15.93   0.000     171.7373    220.0835
       _cons |  -161.8937    39.4527    -4.10   0.000     -239.429   -84.35833
------------------------------------------------------------------------------
. est store y2
. tobit y2 x, ll(0)
Tobit regression                                Number of obs     =        450
                                                LR chi2(1)        =     211.53
                                                Prob > chi2       =     0.0000
Log likelihood =  -2794.467                     Pseudo R2         =     0.0365
------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   232.3739   14.69983    15.81   0.000     203.4849    261.2629
       _cons |  -303.0343   47.96677    -6.32   0.000    -397.3016   -208.7671
-------------+----------------------------------------------------------------
      /sigma |     305.08   11.30586                      282.8611     327.299
------------------------------------------------------------------------------
            67  left-censored observations at y2 <= 0
           383     uncensored observations
             0 right-censored observations
. est store y2tb
. lrtest y2 y2tb, force
Likelihood-ratio test                                 LR chi2(1)  =    732.14
(Assumption: y2 nested in y2tb)                       Prob > chi2 =    0.0000
```

Ans. we reject the null hypothesis in the test, it means that tobit model should be used instead as we have the censored problem. When data is not available and we instead use zero as the information (however, it does not mean that it is zero but it means that it is not available), we have censored problem. If we ignore and run ols, the model will be biased because we got the wrong information.

(d) Estimate the model (3) for y3i using OLS and Tobit model. Determine the most appropriated model in this case. Why? What is the major problem in this case? What will happen if we ignore the problem? (5 points)

```
. reg y3 x

      Source |       SS           df       MS      Number of obs   =       450
-------------+----------------------------------   F(1, 448)       =     34.84
       Model |  6.8534e+09         1  6.8534e+09   Prob > F        =    0.0000
    Residual |  8.8130e+10       448   196718848   R-squared       =    0.0722
-------------+----------------------------------   Adj R-squared   =    0.0701
       Total |  9.4983e+10       449   211544428   Root MSE        =     14026

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   3740.262   633.6827     5.90   0.000     2494.902    4985.622
       _cons |   -7999.71    2032.54    -3.94   0.000    -11994.21   -4005.212
------------------------------------------------------------------------------

. est store y3

. sum y3

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          y3 |        450    3344.719    14544.57  -466.0042   98951.63

. reg y3 x if y3<=1000

      Source |       SS           df       MS      Number of obs   =       427
-------------+----------------------------------   F(1, 425)       =    200.28
       Model |  15439301.7         1  15439301.7   Prob > F        =    0.0000
    Residual |    32762809       425  77088.9623   R-squared       =    0.3203
-------------+----------------------------------   Adj R-squared   =    0.3187
       Total |  48202110.7       426  113150.495   Root MSE        =    277.65

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   190.2182   13.44109    14.15   0.000     163.7989    216.6375
       _cons |  -196.6214   41.95694    -4.69   0.000    -279.0904   -114.1525
------------------------------------------------------------------------------

. est store y3o

. tobit y3 x, ul(1000) nolog

Tobit regression                                Number of obs     =        450
                                                LR chi2(1)        =     212.42
                                                Prob > chi2       =     0.0000
```

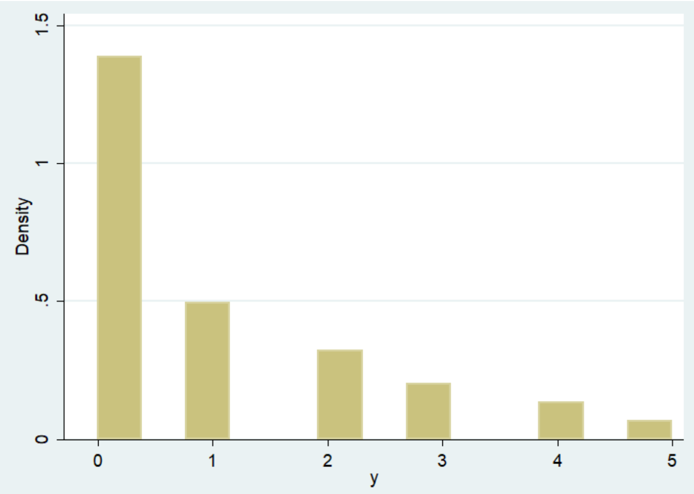
```
Log likelihood = -3056.8296                     Pseudo R2         =     0.0336
------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   219.8752    13.5328    16.25   0.000     193.2798    246.4707
       _cons |  -260.2693   43.08798    -6.04   0.000    -344.9484   -175.5901
-------------+----------------------------------------------------------------
      /sigma |   293.8412   10.15613                      273.8817    313.8006
------------------------------------------------------------------------------
             0  left-censored observations
           427     uncensored observations
            23 right-censored observations at y3 >= 1000
. est store y3ot
. lrtest y3 y3ot, force
Likelihood-ratio test                                 LR chi2(1)  =   3755.16
(Assumption: y3 nested in y3ot)                       Prob > chi2 =    0.0000
. est table y1 y1t y2 y2tb y3 y3ot, star(.1 .05 .01) stat(N rss ll rmse F chi2 r2
r2_a)
------------------------------------------------------------------------------
    Variable |     y1            y1t            y2            y2tb           y3
-------------+----------------------------------------------------------------
_            |
           x |  166.2833***                 195.91036***                 3740.2619***
       _cons | -10.172825                  -161.89367***                -7999.7097***
-------------+----------------------------------------------------------------
eq1          |
           x |                230.09033***
       _cons |               -293.07177***
-------------+----------------------------------------------------------------
sigma        |
       _cons |                  304.355***                 305.08005***
-------------+----------------------------------------------------------------
model        |
           x |                                             232.37386***
       _cons |                                            -303.03434***
-------------+----------------------------------------------------------------
Statistics   |
           N |       367           367           450           450           450
         rss |   24184455                   33204659                     8.813e+10
          ll | -2556.8405    -2523.2473    -3160.5368     -2794.467    -4934.4095
        rmse |  257.40781                  272.24537                      14025.65
           F |  139.84058                  253.68578                     34.838576
```

```
        chi2 |                  106.8811                        211.52829
          r2 |  .27699948                       .36153758                        .07215367
        r2_a |  .27501866                       .36011244                        .07008259
-----------------------------------------------------------------------------------------------
                                                            legend: * p<.1; ** p<.05; *** p<.01
```

```
------------------------------
    Variable |     y3ot
-------------+----------------
_            |
           x |
       _cons |
-------------+----------------
eq1          |
           x |
       _cons |
-------------+----------------
sigma        |
       _cons |  293.84115***
-------------+----------------
model        |
           x |  219.87524***
       _cons | -260.26928***
-------------+----------------
Statistics   |
           N |        450
         rss |
          ll | -3056.8296
        rmse |
           F |
        chi2 |   212.4238
          r2 |
        r2_a |
------------------------------
legend: * p<.1; ** p<.05; *** p<.01
```

ans. from the test we reject the null hypothesis which means tobit model should be used. we censored at 1000 for outliers, the model will be biased if we include all the actual outliers as the result will bend to the side of outliers.

3.(a) Estimate models for Yi assuming that the model is traditional linear regression model. Create histogram for Yi. Determine whether there is limitation of dependent variable in this case. If yes, what type of limitation is it? (3 points)




```
. reg y x1 x2 x3 x4

      Source |       SS           df       MS      Number of obs   =       195
-------------+----------------------------------   F(4, 190)       =      5.10
       Model |  34.7103145         4  8.67757863   Prob > F        =    0.0006
    Residual |  323.289685       190  1.70152466   R-squared       =    0.0970
-------------+----------------------------------   Adj R-squared   =    0.0779
       Total |         358       194  1.84536082   Root MSE        =    1.3044

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0782377   .0455503     1.72   0.087    -.0116115     .168087
          x2 |    .135845   .0482593     2.81   0.005     .0406522    .2310378
          x3 |   .1982819   .0696215     2.85   0.005     .0609515    .3356123
          x4 |  -.0411306   .0493929    -0.83   0.406    -.1385595    .0562983
       _cons |   .9282072    .111482     8.33   0.000     .7083058    1.148109
------------------------------------------------------------------------------

. est store ols
```

Ans. although the ols model seems to be significant in overall, the actual distribution may not be normal as can be seen from the histogram. The distribution of dependent variable is likely to be poisson.

(b) Estimate models for Yi assuming that the probability functions follow Poisson probability distribution. Perform GOF test and determine whether Poisson is appropriated in this case. Interpret the estimated result (sign and meaning (in term of incidence-rate ratios), overall test, individual test, pseudo R 2 , marginal effects). (5 points)

```
. poisson y x1 x2 x3 x4, nolog

Poisson regression                              Number of obs     =        195
                                                LR chi2(4)        =      35.29
```

```
                                                Prob > chi2       =     0.0000
Log likelihood = -276.58677                     Pseudo R2         =     0.0600
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0786569   .0342863     2.29   0.022      .011457    .1458569
          x2 |   .1355416   .0365243     3.71   0.000     .0639553    .2071279
          x3 |   .2110939   .0556809     3.79   0.000     .1019613    .3202265
          x4 |  -.0411834   .0379486    -1.09   0.278    -.1155614    .0331946
       _cons |  -.1600951   .0940253    -1.70   0.089    -.3443813    .0241911
------------------------------------------------------------------------------
. estat gof
         Deviance goodness-of-fit =  321.5064
         Prob > chi2(190)         =    0.0000

         Pearson goodness-of-fit  =   336.287
         Prob > chi2(190)         =    0.0000
. poisson y x1 x2 x3 x4, ir nolog

Poisson regression                              Number of obs     =        195
                                                LR chi2(4)        =      35.29
                                                Prob > chi2       =     0.0000
Log likelihood = -276.58677                     Pseudo R2         =     0.0600

------------------------------------------------------------------------------
           y |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.081833   .0370921     2.29   0.022     1.011523    1.157031
          x2 |   1.145157   .0418261     3.71   0.000     1.066045     1.23014
          x3 |   1.235028   .0687675     3.79   0.000     1.107341     1.37744
          x4 |   .9596531   .0364175    -1.09   0.278     .8908659    1.033752
       _cons |   .8520628   .0801155    -1.70   0.089     .7086587    1.024486
------------------------------------------------------------------------------
. est store poi
. mfx
Marginal effects after poisson
      y  = Predicted number of events (predict)
         =  .91168981
```

```
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0717107       .0309    2.32   0.020   .011153  .132268  -.269757
      x2 |   .1235719      .03234    3.82   0.000   .060179  .186965   .878272
      x3 |   .1924522      .04891    3.93   0.000   .096583  .288322  -.297812
      x4 |  -.0375465       .0345   -1.09   0.276  -.105169  .030076  -.793566
------------------------------------------------------------------------------
```

Ans. in this case we assume passion distribution and perform the test of appropriateness, it turns out that we reject null hypothesis which means poisson model should not be used. The sign and meaning: x4 has negative impact on dependent variable according to IRR<1 and mfx<0. Overall test is significant, pseudo r2 is 0.06, not quite fit the data (we need higher pseudo r2 for better goodness of fit). Individual test for x4 is insignificant, others are significant.

(c) Estimate models for Yi assuming that the probability functions follow Negative Binomial probability distribution. Determine whether Negative Binomial regression model is appropriated in this case. Interpret your estimated result (sign and meaning (in term of incidence-rate ratios), overall test, individual test, pseudo R2 , marginal effects).

```
. nbreg y x1 x2 x3 x4, nolog

Negative binomial regression                      Number of obs     =        195
                                                  LR chi2(4)        =      18.62
Dispersion     = mean                             Prob > chi2       =     0.0009
Log likelihood = -260.81745                       Pseudo R2         =     0.0345
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .1048917   .0532951     1.97   0.049     .0004352    .2093482
          x2 |   .1525739   .0528768     2.89   0.004     .0489372    .2562105
          x3 |   .2032611   .0725929     2.80   0.005     .0609817    .3455405
          x4 |  -.0428056   .0512043    -0.84   0.403    -.1431641    .0575529
       _cons |  -.1783695   .1248751    -1.43   0.153    -.4231201    .0663812
-------------+----------------------------------------------------------------
    /lnalpha |  -.1410752   .2857036                     -.7010441    .4188936
-------------+----------------------------------------------------------------
       alpha |    .868424   .2481119                      .4960671    1.520279
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =   31.54 Prob>=chibar2 = 0.000

. nbreg y x1 x2 x3 x4, ir nolog

Negative binomial regression                      Number of obs     =        195
                                                  LR chi2(4)        =      18.62
Dispersion     = mean                             Prob > chi2       =     0.0009
```

```
Log likelihood = -260.81745                        Pseudo R2         =     0.0345

------------------------------------------------------------------------------
           y |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |    1.11059    .059189     1.97   0.049     1.000435    1.232874
          x2 |   1.164828   .0615924     2.89   0.004     1.050154    1.292025
          x3 |   1.225392   .0889547     2.80   0.005     1.062879    1.412753
          x4 |   .9580976   .0490587    -0.84   0.403     .8666118    1.059241
       _cons |   .8366333   .1044746    -1.43   0.153         .655    1.068634
-------------+----------------------------------------------------------------
    /lnalpha |  -.1410752   .2857036                     -.7010441    .4188936
-------------+----------------------------------------------------------------
       alpha |    .868424   .2481119                      .4960671    1.520279
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =   31.54 Prob>=chibar2 = 0.000
. est store nb
. mfx
Marginal effects after nbreg
      y  = Predicted number of events (predict)
         =  .90553634
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0949833      .04814    1.97   0.049   .000624  .189343  -.269757
      x2 |   .1381612      .04773    2.89   0.004   .044618  .231704   .878272
      x3 |   .1840603      .06547    2.81   0.005   .055747  .312374  -.297812
      x4 |  -.0387621      .04637   -0.84   0.403  -.129637  .052113  -.793566
------------------------------------------------------------------------------
```

Ans. from LR test, the null hypothesis is rejected so alpha exists. Thus, the distribution of y follows negative binomial. So negative binomial regression model should be used in this case. The pseudo r2 is 0.0345, not quite fit the data. From mfx and IRR, x4 has negative impact on y while other have positive. Individual tests are significant except for x4. Overall test is significant

(d) Estimate models for Yi assuming that the model is Zero Inflated Poisson (x1i, x2i, and x3i are independent variables in Poisson model and x4i is independent variable in Inflated (Logit) model). Interpret your estimated result. Determine which model (Linear regression model, Poisson, Negative Binomial, or ZIP) is the most appropriated model in this case? Why? (provide the tests)

```
. zip y x1 x2 x3, inflate(x4) vuong nolog
Zero-inflated Poisson regression                   Number of obs     =        195
                                                   Nonzero obs       =         91
```

```
                                                  Zero obs          =        104

Inflation model = logit                           LR chi2(3)        =      13.67
Log likelihood  = -259.6744                       Prob > chi2       =     0.0034
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
          x1 |   .0882506   .0438924     2.01   0.044     .0022231    .1742781
          x2 |   .1174633   .0406381     2.89   0.004     .0378141    .1971125
          x3 |   .1481343   .0569708     2.60   0.009     .0364735     .259795
       _cons |   .3473148   .1169671     2.97   0.003     .1180636     .576566
-------------+----------------------------------------------------------------
inflate      |
          x4 |   .0307972   .1020081     0.30   0.763    -.1691351    .2307294
       _cons |  -.4919609   .2559837    -1.92   0.055    -.9936799     .009758
------------------------------------------------------------------------------
Vuong test of zip vs. standard Poisson:            z =     2.61  Pr>z = 0.0045
. mfx
Marginal effects after zip
      y  = Predicted number of events (predict)
         =  .91818212
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1 |   .0810301      .04001    2.03   0.043   .002605  .159455  -.269757
      x2 |   .1078527      .03659    2.95   0.003   .036132  .179574   .878272
      x3 |   .1360142      .05122    2.66   0.008    .03562  .236409  -.297812
      x4 |  -.0105671      .03513   -0.30   0.764  -.079421  .058287  -.793566
------------------------------------------------------------------------------
. est store zip
. est table ols poi nb zip, star(.1 .05 .01) stat(N ll chi2 chi2_c voung)
------------------------------------------------------------------------------
    Variable |      ols             poi              nb              zip
-------------+----------------------------------------------------------------
_            |
          x1 |  .07823774*
          x2 |    .135845***
```

```
          x3 |  .19828193***
          x4 | -.04113061
       _cons |  .92820718***
-------------+--------------------------------------------------------------
y            |
          x1 |                .07865693**     .10489172**     .08825059**
          x2 |                 .1355416***    .15257385***    .11746332***
          x3 |                 .2110939***    .20326108***    .14813427***
          x4 |               -.04118338      -.04280564
       _cons |               -.16009506*     -.17836946        .3473148***
-------------+--------------------------------------------------------------
lnalpha      |
       _cons |                               -.14107523
-------------+--------------------------------------------------------------
inflate      |
          x4 |                                                .03079715
       _cons |                                               -.49196094*
-------------+--------------------------------------------------------------
Statistics   |
           N |       195             195             195             195
          ll | -325.98406      -276.58677      -260.81745      -259.67438
        chi2 |                 35.286302        18.62358       13.668238
      chi2_c |                                 31.538645
       voung |
----------------------------------------------------------------------------
                                           legend: * p<.1; ** p<.05; *** p<.01
```

Ans. from the test we reject null; thus, zip should be used. The overall test is significant, individual tests are sig. except for x4. From mfx and IRR, x4 has negative impact on y while other have positive.

From gof, we reject null; thus, poisson is inappropriated.

From test of alpha, the null is rejected; thus negative binomial should be used instead of poisson.

From vuong test, null is rejected; ZIP should be used instead of poisson.

As can be seen from histogram, there are many zeros of dependent variables; thus, ZIP is the most appropriated in this case.

4.(a) Test whether the series yt and xt are stationary series and determine their order of integration. Give explanation of the process of the test. Why is testing equation important? Why is stationary property important to time series analysis?

```
. tsset t

        time variable:  t, 1 to 500
```

```
               delta:  1 unit

. dfuller y, trend lags(1) regress

Augmented Dickey-Fuller test for unit root         Number of obs   =       498

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)            -15.064            -3.980            -3.420            -3.130
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

------------------------------------------------------------------------------
D.y          |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         L1. |  -.9766545   .0648316   -15.06   0.000    -1.104034   -.8492748
         LD. |  -.0591717   .0449902    -1.32   0.189    -.1475675    .0292241
      _trend |    .001712   .0021085     0.81   0.417    -.0024307    .0058547
       _cons |   .8158665   .6108936     1.34   0.182    -.3844037    2.016137
```

```
. reg y x

      Source |       SS           df       MS      Number of obs   =       500
-------------+----------------------------------   F(1, 498)       =  57414.60
       Model |  22599.9131         1  22599.9131   Prob > F        =    0.0000
    Residual |  196.026037       498  .39362658    R-squared       =    0.9914
-------------+----------------------------------   Adj R-squared   =    0.9914
       Total |  22795.9391       499  45.6832448   Root MSE        =     .6274

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |    .697591   .0029113   239.61   0.000      .691871     .703311
       _cons |   .5169709   .0282382    18.31   0.000     .4614903    .5724515
------------------------------------------------------------------------------

. dwstat

Durbin-Watson d-statistic(  2,   500) =  1.699397
```

Ans. from the test, we reject the null; thus, the series is already stationary and the order of integration is zero. The process is that firstly we test the unit root and in this case we reject the null hypothesis of there is unit root; thus, there is no unit root and series is stationary. If we did not reject null, we must test futher by droping other terms in the equation. Before the analysis, we must make

```
sure that the series is stationary. If not, it will lead to spurious problem which
is the false relationship of the variables. Moreover, from DWstat and r2, r2 is not
greateer than DW; thus, there is no spurious problem, supporting the fact that the
series is already stationary.
```

(b) Estimate Autoregressive Integrated Moving Average (ARIMA(p,d,q)) model for yt – determine the most appropriated order for p, d, and q using SBIC given the maximum lag equals 4. Make dynamic forecast for period time = 501 to 505.

```
. est table a10*, star(.1 .05 .01) stat(N ll chi2 bic aic)

------------------------------------------------------------------------------

    Variable |     a101             a102             a103             a104

-------------+----------------------------------------------------------------

y            |

       _cons |  1.2783863***     1.281122***     1.2791531***     1.2796143***

-------------+----------------------------------------------------------------

ARMA         |

          ar |

         L1. | -.85046074***    .81947079***    -.85118368***    -.80777998***

             |

          ma |

         L1. |  .80613284***   -.85871207***     .81662977***     .77333988***

         L2. |                  .07330406        .02669094        .02628038

         L3. |                                   .01786564        .04909843

         L4. |                                                    .04871373

-------------+----------------------------------------------------------------

sigma        |

       _cons |   6.728622***    6.7288968***     6.7268756***     6.7195834***

-------------+----------------------------------------------------------------

Statistics   |

           N |        500              500              500              500

          ll | -1662.6636       -1662.6768       -1662.5349       -1661.9889

        chi2 |  50.079218        23.491264        50.136878        41.183747

         bic |  3350.1855        3356.4266        3362.3575        3367.4801

         aic |  3333.3271        3335.3536        3337.0699        3337.9778

------------------------------------------------------------------------------

                                          legend: * p<.1; ** p<.05; *** p<.01

. est table a20*, star(.1 .05 .01) stat(N ll chi2 bic aic)

------------------------------------------------------------------------------

    Variable |     a201             a202             a203             a204

-------------+----------------------------------------------------------------

y            |
```

```
       _cons |   1.281108***     1.280767***    1.2807996***    1.2807296***
-------------+---------------------------------------------------------------
ARMA         |
          ar |
         L1. |  .72953902***    .03547029      -.0062879      -.01786123
         L2. |  .07788416       .72197954***    .72251715***    .68062226**
             |
          ma |
         L1. | -.77112068***   -.05929498      -.03103008      -.01852434
         L2. |                 -.65277107**     -.6551612**    -.62767115**
         L3. |                                  .0245666        .0231331
         L4. |                                                  .0257644
-------------+---------------------------------------------------------------
sigma        |
       _cons |   6.728009***    6.7172995***    6.7157424***    6.7140346***
-------------+---------------------------------------------------------------
Statistics   |
           N |        500             500             500             500
          ll | -1662.6167      -1661.8255      -1661.7079      -1661.5797
        chi2 |  20.216544       20.405714       21.513034        20.14289
         bic |  3356.3064       3360.9386        3366.918       3372.8764
         aic |  3335.2334        3335.651       3337.4157       3339.1595
-----------------------------------------------------------------------------
                                        legend: * p<.1; ** p<.05; *** p<.01
. est table a30*, star(.1 .05 .01) stat(N ll chi2 bic aic)
-----------------------------------------------------------------------------
    Variable |     a301            a302            a303            a304
-------------+---------------------------------------------------------------
y            |
       _cons |  1.2791645***    1.2808262***      1.29709***    1.2970408***
-------------+---------------------------------------------------------------
ARMA         |
          ar |
         L1. | -.84252296**    -.03735992       1.0737121***    1.0727138***
         L2. |  .02986665       .72407495***    .63880969**     .64085496
         L3. |  .01992436       .02515244      -.82793128***   -.82903352***
             |
          ma |
```

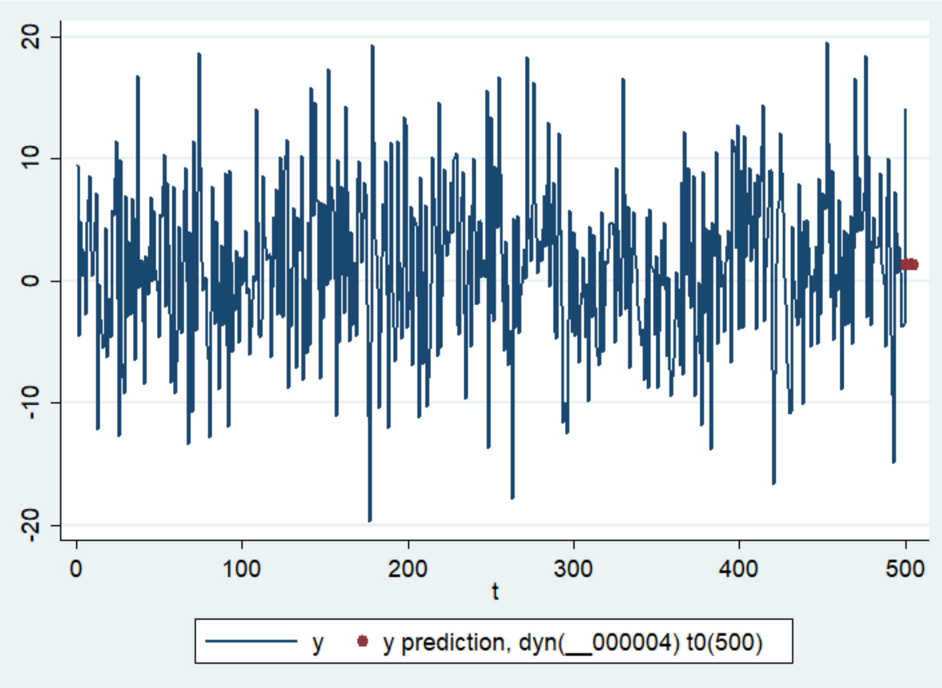
```
         L1. |  .80950138**     .00102216      -1.1331384      -1.1330084
         L2. |                -.65744142**    -.53811447      -.53964878
         L3. |                                 .79692477       .79950405
         L4. |                                                -.00126582
-------------+-----------------------------------------------------------------
sigma        |
       _cons |  6.7268367***    6.7158106***    6.6315795       6.6315672
-------------+-----------------------------------------------------------------
Statistics   |
           N |        500             500             500             500
          ll | -1662.5223      -1661.7165      -1657.4839      -1657.4836
        chi2 |  51.386274       22.059521       48896.509       49072.294
         bic |  3362.3322       3366.9352       3358.4701        3364.684
         aic |  3337.0446        3337.433       3328.9678       3330.9671
-------------------------------------------------------------------------------
                                        legend: * p<.1; ** p<.05; *** p<.01
. est table a40*, star(.1 .05 .01) stat(N ll chi2 bic aic)
-------------------------------------------------------------------------------
    Variable |     a401            a402            a403            a404
-------------+-----------------------------------------------------------------
y            |
       _cons |  1.2798755***     1.281003***    1.2970437***    1.2807299***
-------------+-----------------------------------------------------------------
ARMA         |
          ar |
         L1. | -.76100779**    -.05304826       1.0711477***   -.35555204
         L2. |  .03093655       .64401144*        .642528      -.19846794
         L3. |  .05106124       .02542107      -.8280189***    .30260284
         L4. |  .05080673       .02793854      -.00130282       .6403234**
             |
          ma |
         L1. |  .72690001**     .01626353      -1.1314364***    .32593669
         L2. |                -.59171489      -.54141657       .24606498
         L3. |                                 .79863282***   -.28948366
         L4. |                                                -.55391935*
-------------+-----------------------------------------------------------------
sigma        |
       _cons |  6.7192073***    6.7138919***    6.6315723       6.7074171***
```

```
-------------+----------------------------------------------------------------
Statistics   |
           N |       500             500             500             500
          ll | -1661.9674      -1661.5813      -1657.4836      -1661.1072
        chi2 |  34.131091       19.371191       49155.273       31.137439
         bic |   3367.437       3372.8795        3364.684       3384.3604
         aic |  3337.9347       3339.1627       3330.9671       3342.2143
------------------------------------------------------------------------------
                                          legend: * p<.1; ** p<.05; *** p<.01
```

ans. Given max lags equal to 4, the most appropriated orders for(p,q,d) are (1,0,1) as the model has lowest BIC.

```
. arima y, arima(1,0,1) nolog

ARIMA regression

Sample:  1 - 500                                Number of obs     =        500
                                                Wald chi2(2)      =      50.08
Log likelihood = -1662.664                      Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |                 OPG
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
       _cons |   1.278386   .2956659     4.32   0.000     .6988917    1.857881
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |  -.8504607   .1673848    -5.08   0.000    -1.178529   -.5223925
             |
          ma |
         L1. |   .8061328   .1890731     4.26   0.000     .4355564    1.176709
-------------+----------------------------------------------------------------
      /sigma |   6.728622   .2147767    31.33   0.000     6.307667    7.149577
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
      confidence interval is truncated at zero.

. set obs 505
number of observations (_N) was 500, now 505

. replace t=_n
```

```
(5 real changes made)

. predict yhat, y dynamic(.) t0(500)

Note: beginning dynamic predictions in period         3.

(499 missing values generated)

. twoway (line y t, sort) (scatter yhat t, sort)
```



(c) Estimate model (5) using OLS by employing yt as dependent variable and xt as explanatory variable, determine whether ARCH-effect significantly occurs, and state null hypothesis of the test. Why the test is classified as LM test?

```
. reg y x

      Source |       SS           df       MS      Number of obs   =       500
-------------+----------------------------------   F(1, 498)       =  57414.60
       Model |  22599.9131         1  22599.9131   Prob > F        =    0.0000
    Residual |  196.026037       498   .39362658   R-squared       =    0.9914
-------------+----------------------------------   Adj R-squared   =    0.9914
       Total |  22795.9391       499  45.6832448   Root MSE        =     .6274

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |    .697591   .0029113   239.61   0.000      .691871     .703311
       _cons |   .5169709   .0282382    18.31   0.000     .4614903    .5724515
------------------------------------------------------------------------------

. estat archlm

LM test for autoregressive conditional heteroskedasticity (ARCH)

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

lags(p)	chi2	df	Prob > chi2
1	38.569	1	0.0000

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Ans. the null hypothesis is no arch effects. In this case we reject; thus, there is significant arch problem. Here for ols, it is a restricted model as we assume no arch effect in ols. So, the test will be LM test as the model is restricted.

(d) Estimate GARCH(p,q) for yt using xt as explanatory variable for mean equation (model (5) and (6)) – determine the most appropriated order p and q for variance equation using SBIC given the maximum lag equals to 2. Then, predict the variance of yt using the estimated result of GARCH(p,q) model with the most appropriated lag.

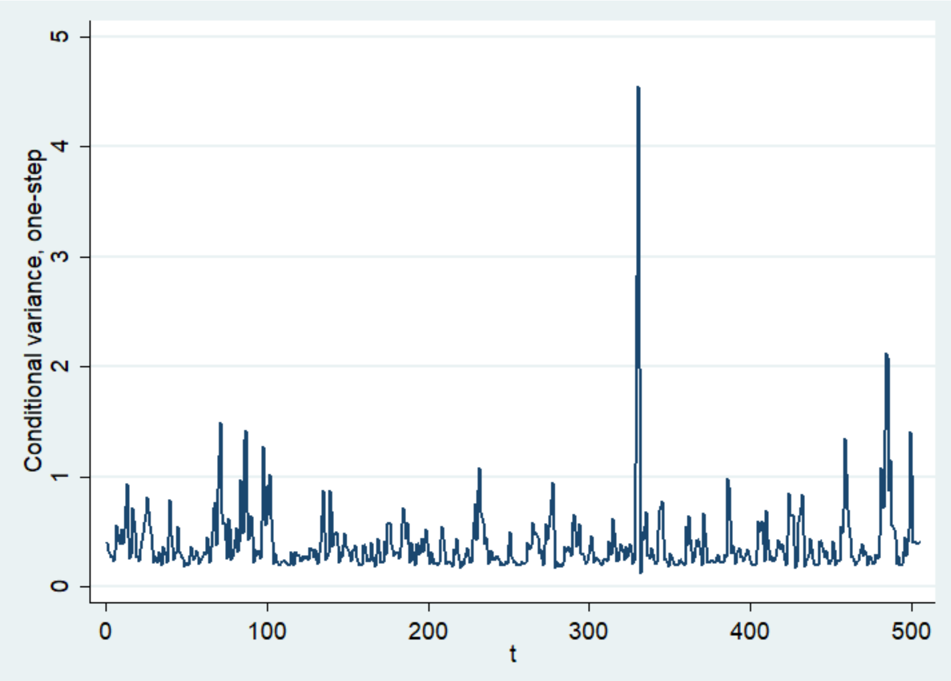
```
. est table g*, star(.1 .05 .01) stat(N ll chi2 bic aic)
```

Variable	g11	g12	g21	g22
y				
x	.69775812***	.69774139***	.69774644***	.69762642***
_cons	.51842966***	.51839207***	.51840685***	.51595866***
ARCH				
arch				
L1.	.36223733***	.36055372***	.36091784***	.37249344***
L2.		.015322		.30920474**
garch				
L1.	.31538419***	.28751012	.32622986*	-.39550084
L2.			-.0097377	.10865385
_cons	.128928***	.13460002*	.12900729***	.24612001***
Statistics				
N	500	500	500	500
ll	-443.6784	-443.67152	-443.67318	-442.8689
chi2	72792.196	72882.484	72850.286	74372.71
bic	918.42983	924.63069	924.63401	929.24006
aic	897.35679	899.34304	899.34636	899.7378

legend: * p<.1; ** p<.05; *** p<.01

```
. predict sigmahat, v
```

```
. line sigmahat t
```




```
Ans. according to lowest BIC, garch(1,1) is the most appropriated one.
```

(e) Among these three models, ARCH(1), GARCH(1,1), or EGARCH(1,1,1), which model is the most appropriated in this case? Why? What are the differences among ARCH, GARCH, and EGARCH?

```
. arch y x, egarch(1/1) garch(1/1) arch(1/1) nolog

ARCH family regression

Sample: 1 - 500                                 Number of obs     =        500
Distribution: Gaussian                          Wald chi2(1)      =   67004.12
Log likelihood = -447.1185                      Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |                 OPG
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
           x |   .6979221   .0026962   258.85   0.000     .6926376    .7032066
       _cons |   .5116238   .0250266    20.44   0.000     .4625726     .560675
-------------+----------------------------------------------------------------
ARCH         |
      egarch |
         L1. |   .3619726   .1238794     2.92   0.003     .1191734    .6047719
             |
        arch |
         L1. |   .6726571   .0925512     7.27   0.000     .4912601    .8540541
```

```
             |
       garch |
         L1. |  -.0060473   .0106304    -0.57   0.569    -.0268825    .0147878
             |
       _cons |  -.9232664   .1570025    -5.88   0.000    -1.230986   -.6155472
------------------------------------------------------------------------------

. est store eg111

. est table eg111, star(.1 .05 .01) stat(N ll chi2 bic aic)

------------------------------
    Variable |     eg111
-------------+----------------
y            |
           x |  .69792214***
       _cons |  .51162381***
-------------+----------------
ARCH         |
      egarch |
         L1. |  .36197264***
             |
        arch |
         L1. |  .67265712***
             |
       garch |
         L1. | -.00604733
             |
       _cons | -.92326637***
-------------+----------------
Statistics   |
           N |        500
          ll |  -447.1185
        chi2 |  67004.123
         bic |  931.52464
         aic |  906.23699
------------------------------
legend: * p<.1; ** p<.05; *** p<.01
```

Ans. from lowest BIC, garch(1,1) is most appropriated as it has the lowest value of BIC. Arch model is a process of weighted sum of lags of error terms or white noises. Garch is a model where we model the lags terms of variance. For egarch, it is when we have asymmetric model, the garch graph is not linear but kinked. In egarch, the exponential term is included to make the graph more smooth. Moreover, in practice, we need high frequency data to detect the both arch and garch effect. But for garch effects, the frequency must be like daily data.

5.(a) Estimate VARs models using yt and xt as endogenous variables and determine the most appropriated lags models using SBIC

```
. varsoc y x

   Selection-order criteria

   Sample:  5 - 500                              Number of obs      =       496

  +---------------------------------------------------------------------------+

  |lag |    LL      LR      df    p       FPE       AIC      HQIC      SBIC    |

  |----+----------------------------------------------------------------------|

  |  0 | -1366.52                      .854238   5.51821   5.52487   5.53517  |

  |  1 | -1320.19  92.648*   4  0.000  .720214*  5.34755*  5.36752*  5.39843* |

  |  2 | -1319.33  1.7247    4  0.786  .729385    5.3602   5.39349   5.44501  |

  |  3 | -1318.25  2.1559    4  0.707  .738031   5.37198   5.41859   5.49072  |

  |  4 | -1316.13  4.2352    4  0.375  .743657   5.37957   5.43949   5.53223  |

  +---------------------------------------------------------------------------+

   Endogenous:  y x

    Exogenous:  _cons
```

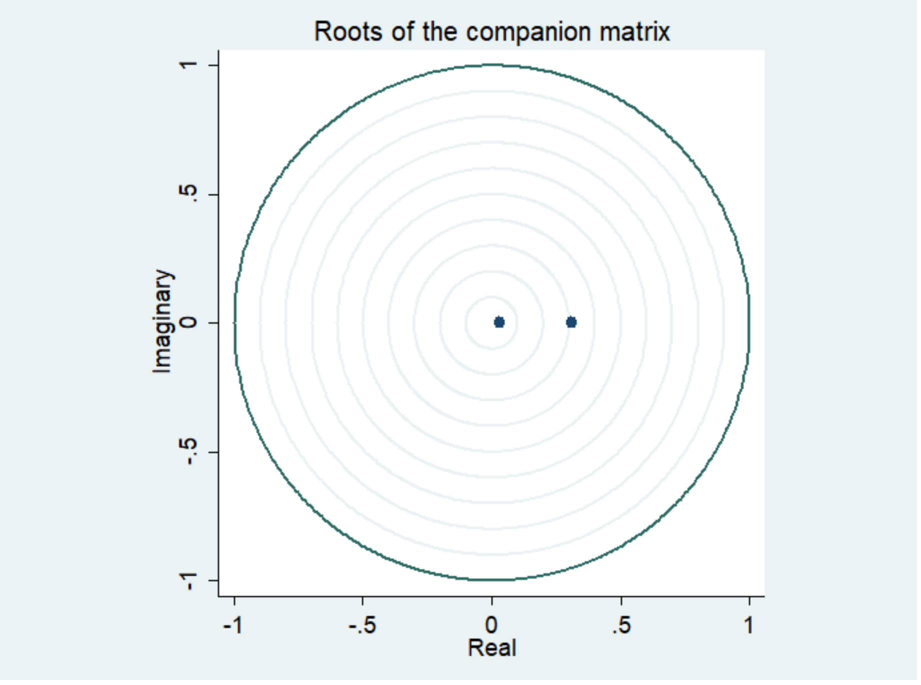
Ans. according to lowest BIC, the most appropriated lag order is one.

(b) Perform stability test and Granger exogeneity test. Determine whether assumptions of VARs are satisfied, explain your evaluation criteria. If the stability assumption is unsatisfied, what will happen?

```
. var y x, lag(1/1)

Vector autoregression

Sample:  2 - 500                                 Number of obs      =       499

Log likelihood =   -1330.31                      AIC                =  5.355953

FPE            =   .7262936                      HQIC               =  5.375831

Det(Sigma_ml)  =   .7090358                      SBIC               =  5.406605

Equation           Parms      RMSE     R-sq      chi2     P>chi2

----------------------------------------------------------------

y                     3     .994284   0.0148   7.496551   0.0236
```

```
x                    3     .998231   0.0871   47.60214   0.0000
----------------------------------------------------------------
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
           y |
         L1. |   .1368967    .050075     2.73   0.006     .0387515     .235042
             |
           x |
         L1. |   .0538896   .0479744     1.12   0.261    -.0401384    .1479176
             |
       _cons |   .3381557   .0544482     6.21   0.000     .2314393    .4448722
-------------+----------------------------------------------------------------
x            |
           y |
         L1. |   .3397947   .0502738     6.76   0.000     .2412598    .4383296
             |
           x |
         L1. |   .2089573   .0481648     4.34   0.000      .114556    .3033586
             |
       _cons |   .1240954   .0546643     2.27   0.023     .0169553    .2312355
------------------------------------------------------------------------------
. varstable, graph
   Eigenvalue stability condition
  +----------------------------------------+
  |        Eigenvalue        |   Modulus   |
  |--------------------------+-------------|
  |   .3129612               |   .312961   |
```

```
  |   .0328928              |   .032893   |
  +----------------------------------------+
   All the eigenvalues lie inside the unit circle.
   VAR satisfies stability condition.
```




```
. vargranger

   Granger causality Wald tests
  +------------------------------------------------------------------+
  |          Equation           Excluded |   chi2     df Prob > chi2 |
  |--------------------------------------+---------------------------|
  |                 y                  x |  1.2618     1    0.261    |
  |                 y                ALL |  1.2618     1    0.261    |
  |--------------------------------------+---------------------------|
  |                 x                  y |  45.682     1    0.000    |
  |                 x                ALL |  45.682     1    0.000    |
  +------------------------------------------------------------------+
```

Ans. from stability test, all eigenvalues lie within unit circle; stability assumption is satisfied and the system is stable. According to granger test, the variables will cause each other if both null hypotheses are rejected. In this case, only one is rejected; x seems not to be the endogenous variable but y is an endogenous variable. If the system is not stable, the impulse will not collapse into zero when we do IRF.

(c) Perform Impulse response function analysis (irf), Orthogonal impulse response function analysis (oirf), Cumulative impulse response function analysis (coirf), make interpretation of the analysis, and determine which variable has more impact (using Cholesky order – yt xt).

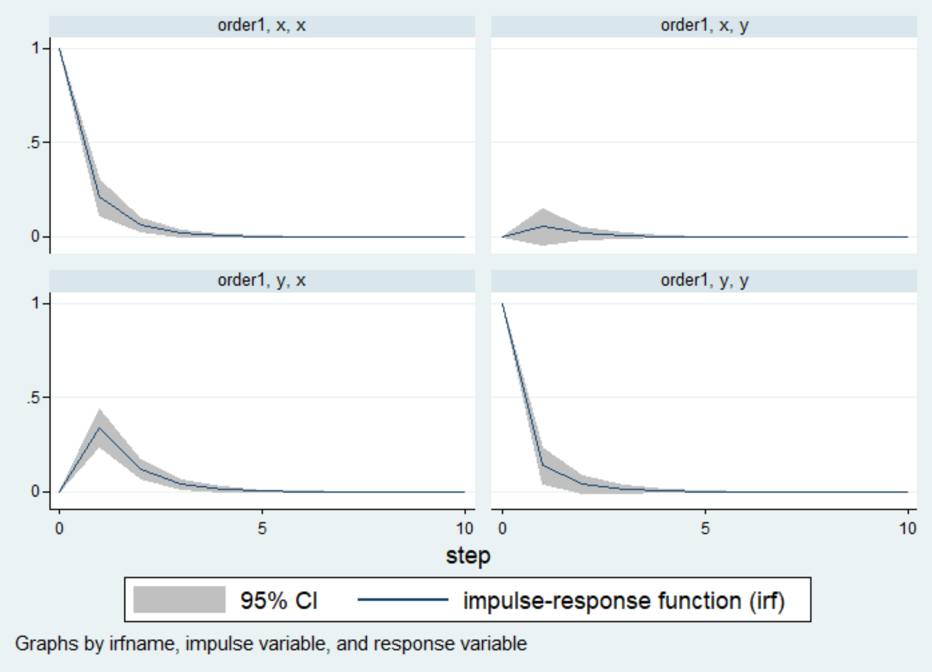
```
. irf create order1, o(y x) step(10) set(irf_1)

(file irf_1.irf created)

(file irf_1.irf now active)

(file irf_1.irf updated)

. irf table irf, impulse(y x) response(y x)
```

Results from order1

step	(1) irf	(1) Lower	(1) Upper	(2) irf	(2) Lower	(2) Upper
0	1	1	1	0	0	0
1	.136897	.038751	.235042	.339795	.24126	.43833
2	.037052	-.011225	.085329	.117519	.066423	.168615
3	.011405	-.008221	.031032	.037147	.012582	.061712
4	.003563	-.003834	.01096	.011638	.000444	.022831
5	.001115	-.001591	.00382	.003642	-.001081	.008366
6	.000349	-.000621	.001319	.00114	-.000736	.003016
7	.000109	-.000233	.000452	.000357	-.000357	.001071
8	.000034	-.000085	.000154	.000112	-.000152	.000375
9	.000011	-.000031	.000052	.000035	-.00006	.00013
10	3.3e-06	-.000011	.000017	.000011	-.000023	.000045

	(3)	(3)	(3)	(4)	(4)	(4)

```
|  step  |   irf        Lower       Upper     |   irf        Lower       Upper     |
|--------+----------------------------------+----------------------------------|
|0       | 0           0           0         | 1           1           1         |
|1       | .05389      -.040138    .147918   | .208957     .114556     .303359   |
|2       | .018638     -.014291    .051566   | .061975     .024515     .099434   |
|3       | .005891     -.006206    .017989   | .019283     .001614     .036952   |
|4       | .001846     -.002574    .006265   | .006031     -.00155     .013612   |
|5       | .000578     -.00101     .002165   | .001887     -.001149    .004923   |
|6       | .000181     -.000381    .000742   | .000591     -.00057     .001752   |
|7       | .000057     -.000139    .000253   | .000185     -.000245    .000615   |
|8       | .000018     -.00005     .000085   | .000058     -.000098    .000213   |
|9       | 5.5e-06     -.000018    .000029   | .000018     -.000037    .000073   |
|10      | 1.7e-06     -6.2e-06    9.6e-06   | 5.7e-06     -.000014    .000025   |
+-----------------------------------------------------------------------------+
95% lower and upper bounds reported
(1) irfname = order1, impulse = y, and response = y
(2) irfname = order1, impulse = y, and response = x
(3) irfname = order1, impulse = x, and response = y
(4) irfname = order1, impulse = x, and response = x
. irf graph irf, impulse(y x) response(y x)
```




```
. irf table oirf, impulse(y x) response(y x)

                              Results from order1

+-----------------------------------------------------------------------------+
|        |    (1)          (1)          (1)     |    (2)          (2)          (2)     |
|  step  |   oirf        Lower        Upper     |   oirf        Lower        Upper     |
|--------+--------------------------------------+--------------------------------|
|0       | .99129       .929789      1.05279    | -.518579     -.599757     -.437401   |
|1       | .107758      .02092       .194597    | .228474      .139126      .317822    |
|2       | .027064      -.005259     .059387    | .084357      .04536       .123355    |
|3       | .008251      -.005158     .02166     | .026823      .009633      .044014    |
|4       | .002575      -.002519     .007669    | .008409      .000835      .015982    |
|5       | .000806      -.001067     .002678    | .002632      -.000565     .005829    |
|6       | .000252      -.000422     .000926    | .000824      -.000455     .002102    |
|7       | .000079      -.00016      .000318    | .000258      -.000232     .000748    |
|8       | .000025      -.000059     .000108    | .000081      -.000101     .000263    |
|9       | 7.7e-06      -.000021     .000037    | .000025      -.000041     .000091    |
|10      | 2.4e-06      -7.5e-06     .000012    | 7.9e-06      -.000016     .000031    |
+-----------------------------------------------------------------------------+
```

step	(3) oirf	(3) Lower	(3) Upper	(4) oirf	(4) Lower	(4) Upper
0	0	0	0	.849441	.79674	.902142
1	.045776	-.034146	.125698	.177497	.096556	.258438
2	.015832	-.012156	.04382	.052644	.020657	.084631
3	.005004	-.005277	.015285	.01638	.001337	.031423
4	.001568	-.002188	.005323	.005123	-.001324	.01157
5	.000491	-.000858	.001839	.001603	-.000978	.004184
6	.000154	-.000323	.00063	.000502	-.000485	.001489
7	.000048	-.000118	.000215	.000157	-.000208	.000522
8	.000015	-.000043	.000073	.000049	-.000083	.000181
9	4.7e-06	-.000015	.000024	.000015	-.000032	.000062
10	1.5e-06	-5.2e-06	8.2e-06	4.8e-06	-.000012	.000021

95% lower and upper bounds reported

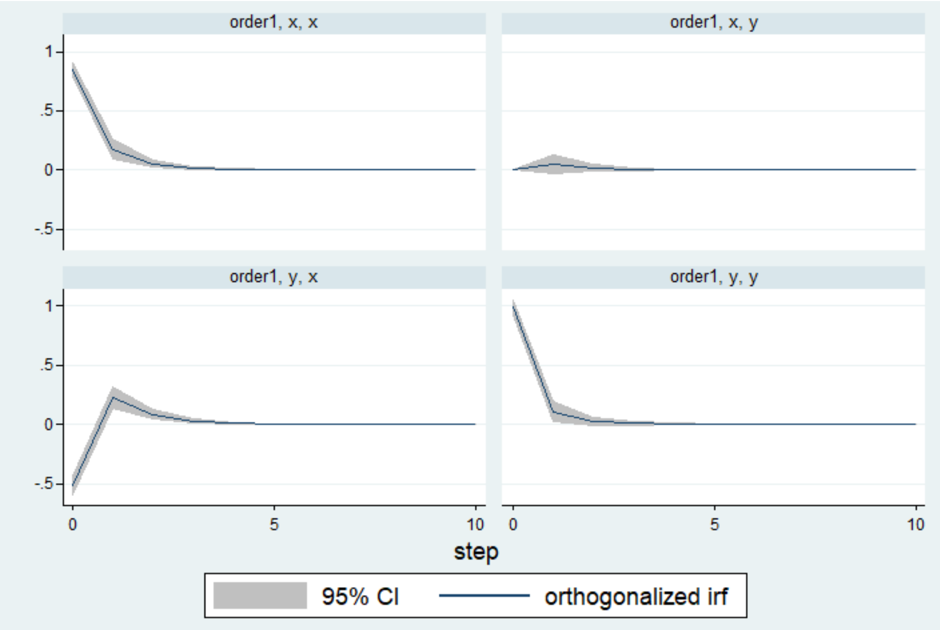
(1) irfname = order1, impulse = y, and response = y

(2) irfname = order1, impulse = y, and response = x

(3) irfname = order1, impulse = x, and response = y

(4) irfname = order1, impulse = x, and response = x

```
. irf graph oirf, impulse(y x) response(y x)
```




```
. irf table coirf, impulse(y x) response(y x)

                              Results from order1

+-----------------------------------------------------------------------------+
|        |    (1)         (1)         (1)     |    (2)         (2)         (2)     |
|  step  |  coirf       Lower       Upper     |  coirf       Lower       Upper     |
|--------+------------------------------------+------------------------------------|
|0       | .99129      .929789     1.05279    | -.518579    -.599757    -.437401   |
|1       | 1.09905     .988841     1.20926    | -.290105    -.416525    -.163684   |
|2       | 1.12611     .997787     1.25444    | -.205747    -.347383    -.064112   |
|3       | 1.13436     .998474     1.27025    | -.178924    -.327699    -.030149   |
|4       | 1.13694     .99813      1.27575    | -.170516    -.322242    -.01879    |
|5       | 1.13774     .997874     1.27761    | -.167884    -.320744    -.015023   |
|6       | 1.138       .997752     1.27824    | -.16706     -.320336    -.013784   |
|7       | 1.13808     .9977       1.27845    | -.166802    -.320226    -.013379   |
|8       | 1.1381      .997681     1.27852    | -.166721    -.320196    -.013247   |
|9       | 1.13811     .997673     1.27854    | -.166696    -.320189    -.013204   |
|10      | 1.13811     .99767      1.27855    | -.166688    -.320187    -.01319    |
+-----------------------------------------------------------------------------+
```

step	(3) coirf	(3) Lower	(3) Upper	(4) coirf	(4) Lower	(4) Upper
0	0	0	0	.849441	.79674	.902142
1	.045776	-.034146	.125698	1.02694	.92452	1.12936
2	.061608	-.045907	.169123	1.07958	.955614	1.20355
3	.066612	-.050942	.184167	1.09596	.963061	1.22886
4	.06818	-.053012	.189371	1.10108	.964611	1.23756
5	.068671	-.053817	.191158	1.10269	.964865	1.24051
6	.068824	-.054118	.191766	1.10319	.964877	1.2415
7	.068872	-.054227	.191972	1.10335	.96486	1.24183
8	.068887	-.054266	.192041	1.1034	.964849	1.24194
9	.068892	-.05428	.192064	1.10341	.964844	1.24198
10	.068893	-.054285	.192072	1.10342	.964842	1.24199

95% lower and upper bounds reported

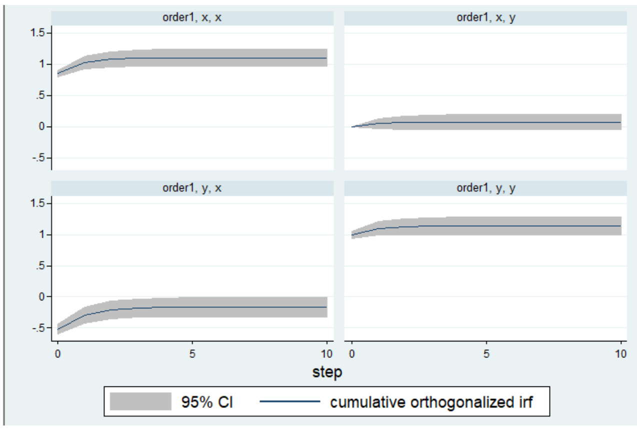
(1) irfname = order1, impulse = y, and response = y

(2) irfname = order1, impulse = y, and response = x

(3) irfname = order1, impulse = x, and response = y

(4) irfname = order1, impulse = x, and response = x

```
. irf graph coirf, impulse(y x) response(y x)
```



Ans. from all the graphs, it seems that there is positive impact. From the coirf graph, it can be seen that the magnitude of impact y on x is larger than x on y; thus, y has more impact on x. Moreover, the impact of x on y mostly insignificant as almost all of them cover zero in the interval.

(d) Perform Forecast error variance decomposition (fevd) and determine variable that has more impact on each endogenous variable

```
. irf table fevd, impulse(y x) response(y x)
```

Results from order1

step	(1) fevd	(1) Lower	(1) Upper	(2) fevd	(2) Lower	(2) Upper
0	0	0	0	0	0	0
1	1	1	1	.27151	.2049	.338121
2	.997897	.990562	1.00523	.298949	.234118	.363779
3	.997648	.989454	1.00584	.302786	.238493	.36708
4	.997623	.989331	1.00591	.303174	.238941	.367406
5	.99762	.989317	1.00592	.303212	.238985	.367438
6	.99762	.989316	1.00592	.303216	.23899	.367441
7	.99762	.989316	1.00592	.303216	.23899	.367442
8	.99762	.989316	1.00592	.303216	.23899	.367442
9	.99762	.989316	1.00592	.303216	.23899	.367442
10	.99762	.989316	1.00592	.303216	.23899	.367442

step	(3) fevd	(3) Lower	(3) Upper	(4) fevd	(4) Lower	(4) Upper
0	0	0	0	0	0	0
1	0	0	0	.72849	.661879	.7951
2	.002103	-.005232	.009438	.701051	.636221	.765882

```
|3        | .002352     -.005842    .010546   | .697214     .63292      .761507   |
|4        | .002377     -.005915    .010669   | .696826     .632594     .761059   |
|5        | .00238      -.005923    .010683   | .696788     .632562     .761015   |
|6        | .00238      -.005924    .010684   | .696784     .632559     .76101    |
|7        | .00238      -.005925    .010684   | .696784     .632558     .76101    |
|8        | .00238      -.005925    .010684   | .696784     .632558     .76101    |
|9        | .00238      -.005925    .010684   | .696784     .632558     .76101    |
|10       | .00238      -.005925    .010684   | .696784     .632558     .76101    |
+-------------------------------------------------------------------------------+
95% lower and upper bounds reported
(1) irfname = order1, impulse = y, and response = y
(2) irfname = order1, impulse = y, and response = x
(3) irfname = order1, impulse = x, and response = y
(4) irfname = order1, impulse = x, and response = x
```

Ans. according to Fevd, y has more impact on x.

(e) Determine whether changing Cholesky order from – “yt xt” to “xt yt” will change the results of irf, oirf, coirf, and fevd.

```
. irf create order1, o(x y) step(10) set(irf_2)
(file irf_2.irf created)
(file irf_2.irf now active)
(file irf_2.irf updated)
. irf table irf, impulse(x y) response(x y)
                                        Results from order1
+-------------------------------------------------------------------------------+
|        |    (1)         (1)         (1)     |    (2)         (2)         (2)    |
|  step  |   irf        Lower       Upper    |   irf        Lower       Upper   |
|--------+-----------------------------------+-----------------------------------|
|0       | 1           1           1         | 0           0           0         |
|1       | .208957     .114556     .303359   | .05389      -.040138    .147918   |
```

2	.061975	.024515	.099434	.018638	-.014291	.051566
3	.019283	.001614	.036952	.005891	-.006206	.017989
4	.006031	-.00155	.013612	.001846	-.002574	.006265
5	.001887	-.001149	.004923	.000578	-.00101	.002165
6	.000591	-.00057	.001752	.000181	-.000381	.000742
7	.000185	-.000245	.000615	.000057	-.000139	.000253
8	.000058	-.000098	.000213	.000018	-.00005	.000085
9	.000018	-.000037	.000073	5.5e-06	-.000018	.000029
10	5.7e-06	-.000014	.000025	1.7e-06	-6.2e-06	9.6e-06

	(3)	(3)	(3)	(4)	(4)	(4)
step	irf	Lower	Upper	irf	Lower	Upper
0	0	0	0	1	1	1
1	.339795	.24126	.43833	.136897	.038751	.235042
2	.117519	.066423	.168615	.037052	-.011225	.085329
3	.037147	.012582	.061712	.011405	-.008221	.031032
4	.011638	.000444	.022831	.003563	-.003834	.01096
5	.003642	-.001081	.008366	.001115	-.001591	.00382
6	.00114	-.000736	.003016	.000349	-.000621	.001319
7	.000357	-.000357	.001071	.000109	-.000233	.000452
8	.000112	-.000152	.000375	.000034	-.000085	.000154
9	.000035	-.00006	.00013	.000011	-.000031	.000052
10	.000011	-.000023	.000045	3.3e-06	-.000011	.000017

95% lower and upper bounds reported

(1) irfname = order1, impulse = x, and response = x

(2) irfname = order1, impulse = x, and response = y

(3) irfname = order1, impulse = y, and response = x

(4) irfname = order1, impulse = y, and response = y

```
. irf table oirf, impulse(x y) response(x y)

                              Results from order1

+-------------------------------------------------------------------------------+
|        |    (1)          (1)          (1)     |    (2)          (2)          (2)     |
|  step  |   oirf         Lower        Upper    |   oirf         Lower        Upper    |
|--------+-----------------------------------+-----------------------------------|
|0       | .995226      .93348       1.05697    | -.516528     -.597385     -.435671  |
|1       | .032446      -.055074     .119966    | -.017079     -.101153     .066995   |
|2       | .000977      -.025324     .027278    | -.00059      -.013299     .01212    |
|3       | 3.8e-06      -.009566     .009573    | -.000028     -.003175     .003119   |
|4       | -8.8e-06     -.003076     .003058    | -3.6e-06     -.00095      .000943   |
|5       | -3.1e-06     -.000966     .000959    | -9.7e-07     -.000295     .000293   |
|6       | -9.7e-07     -.000302     .0003      | -3.0e-07     -.000092     .000092   |
|7       | -3.0e-07     -.000094     .000094    | -9.3e-08     -.000029     .000029   |
|8       | -9.5e-08     -.00003      .000029    | -2.9e-08     -9.0e-06     9.0e-06   |
|9       | -3.0e-08     -9.2e-06     9.2e-06    | -9.1e-09     -2.8e-06     2.8e-06   |
|10      | -9.3e-09     -2.9e-06     2.9e-06    | -2.9e-09     -8.8e-07     8.8e-07   |
+-------------------------------------------------------------------------------+

+-------------------------------------------------------------------------------+
|        |    (3)          (3)          (3)     |    (4)          (4)          (4)     |
|  step  |   oirf         Lower        Upper    |   oirf         Lower        Upper    |
|--------+-----------------------------------+-----------------------------------|
|0       | 0            0            0          | .846082      .79359       .898574   |
|1       | .287494      .202239      .37275     | .115826      .032477      .199175   |
|2       | .099431      .055762      .1431      | .031349      -.009544     .072242   |
|3       | .031429      .010554      .052304    | .00965       -.006967     .026266   |
|4       | .009846      .000356      .019336    | .003015      -.003247     .009276   |
```

5	.003082	-.000919	.007083	.000943	-.001346	.003233
6	.000965	-.000624	.002553	.000295	-.000526	.001116
7	.000302	-.000303	.000906	.000092	-.000198	.000382
8	.000094	-.000129	.000318	.000029	-.000072	.00013
9	.00003	-.000051	.00011	9.0e-06	-.000026	.000044
10	9.3e-06	-.000019	.000038	2.8e-06	-9.1e-06	.000015

95% lower and upper bounds reported

(1) irfname = order1, impulse = x, and response = x

(2) irfname = order1, impulse = x, and response = y

(3) irfname = order1, impulse = y, and response = x

(4) irfname = order1, impulse = y, and response = y

. irf table coirf, impulse(x y) response(x y)

Results from order1

step	(1) coirf	(1) Lower	(1) Upper	(2) coirf	(2) Lower	(2) Upper
0	.995226	.93348	1.05697	-.516528	-.597385	-.435671
1	1.02767	.919409	1.13593	-.533607	-.656826	-.410388
2	1.02865	.911276	1.14602	-.534197	-.667537	-.400856
3	1.02865	.908054	1.14925	-.534225	-.670073	-.398376
4	1.02864	.90693	1.15036	-.534228	-.670828	-.397628
5	1.02864	.906566	1.15071	-.534229	-.671063	-.397395
6	1.02864	.906451	1.15083	-.534229	-.671137	-.397322
7	1.02864	.906415	1.15086	-.53423	-.67116	-.397299
8	1.02864	.906403	1.15087	-.53423	-.671167	-.397292
9	1.02864	.9064	1.15088	-.53423	-.671169	-.39729
10	1.02864	.906399	1.15088	-.53423	-.67117	-.397289

step	(3) coirf	(3) Lower	(3) Upper	(4) coirf	(4) Lower	(4) Upper
0	0	0	0	.846082	.79359	.898574
1	.287494	.202239	.37275	.961908	.859649	1.06417
2	.386925	.269313	.504537	.993257	.859792	1.12672
3	.418354	.286622	.550087	1.00291	.856144	1.14967
4	.428201	.290687	.565714	1.00592	.854101	1.15774
5	.431282	.29153	.571035	1.00687	.853205	1.16052
6	.432247	.291665	.572829	1.00716	.852848	1.16147
7	.432549	.291669	.573428	1.00725	.852713	1.16179
8	.432643	.291659	.573627	1.00728	.852663	1.1619
9	.432673	.291652	.573693	1.00729	.852645	1.16194
10	.432682	.291649	.573715	1.00729	.852639	1.16195

95% lower and upper bounds reported

(1) irfname = order1, impulse = x, and response = x

(2) irfname = order1, impulse = x, and response = y

(3) irfname = order1, impulse = y, and response = x

(4) irfname = order1, impulse = y, and response = y

. irf graph coirf, impulse(x y) response(x y)

. irf table fevd, impulse(x y) response(x y)

Results from order1

step	(1) fevd	(1) Lower	(1) Upper	(2) fevd	(2) Lower	(2) Upper

step	fevd	Lower	Upper	fevd	Lower	Upper
0	0	0	0	0	0	0
1	1	1	1	.27151	.2049	.338121
2	.923055	.879983	.966127	.268068	.201606	.33453
3	.914637	.867335	.961939	.267804	.201324	.334284
4	.913804	.866015	.961594	.267779	.201296	.334262
5	.913723	.86588	.961565	.267777	.201293	.33426
6	.913715	.865866	.961563	.267776	.201292	.33426
7	.913714	.865865	.961563	.267776	.201292	.33426
8	.913714	.865865	.961563	.267776	.201292	.33426
9	.913714	.865865	.961563	.267776	.201292	.33426
10	.913714	.865865	.961563	.267776	.201292	.33426

	(3)	(3)	(3)	(4)	(4)	(4)
step	fevd	Lower	Upper	fevd	Lower	Upper
0	0	0	0	0	0	0
1	0	0	0	.72849	.661879	.7951
2	.076945	.033873	.120017	.731932	.66547	.798394
3	.085363	.038061	.132665	.732196	.665716	.798676
4	.086196	.038406	.133985	.732221	.665738	.798704
5	.086277	.038435	.13412	.732223	.66574	.798707
6	.086285	.038437	.134134	.732224	.66574	.798708
7	.086286	.038437	.134135	.732224	.66574	.798708
8	.086286	.038437	.134135	.732224	.66574	.798708
9	.086286	.038437	.134135	.732224	.66574	.798708
10	.086286	.038437	.134135	.732224	.66574	.798708

95% lower and upper bounds reported

(1) irfname = order1, impulse = x, and response = x

(2) irfname = order1, impulse = x, and response = y

(3) irfname = order1, impulse = y, and response = x

(4) irfname = order1, impulse = y, and response = y

Ans. yes, it will change the result if we change the order. As a result, we need to make sure the order first. In this case, the impact of x on y is larger (previous case is impact on y is larger) when we change the order from (y x) to (x y).