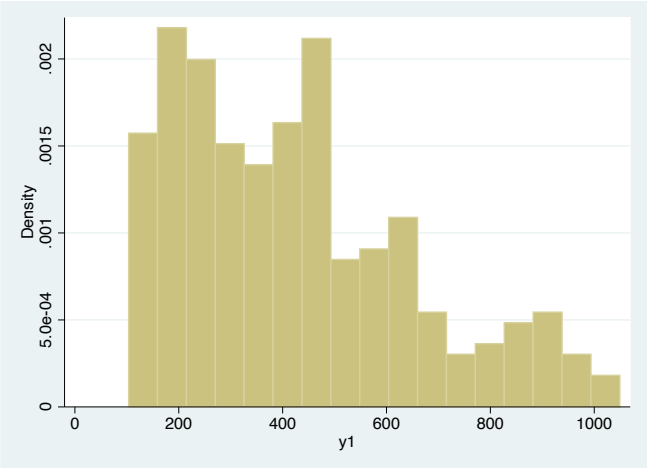
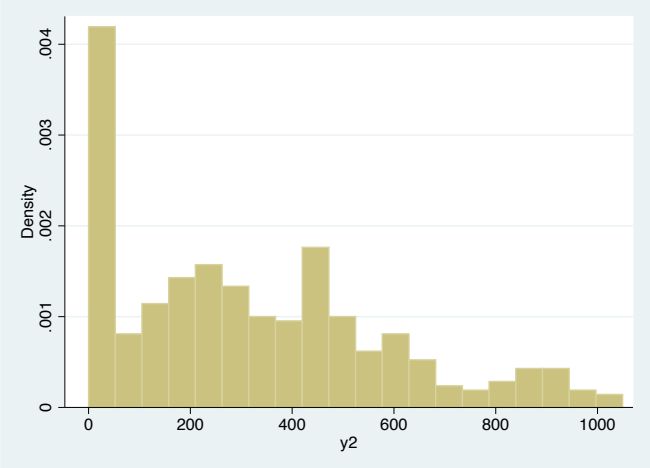
1. Plot histogram of y_{1i}, y_{2i}, y_{3i}, compute descriptive statistics of these three variables, then determine limitation of these three dependent variables.

. histogram y1



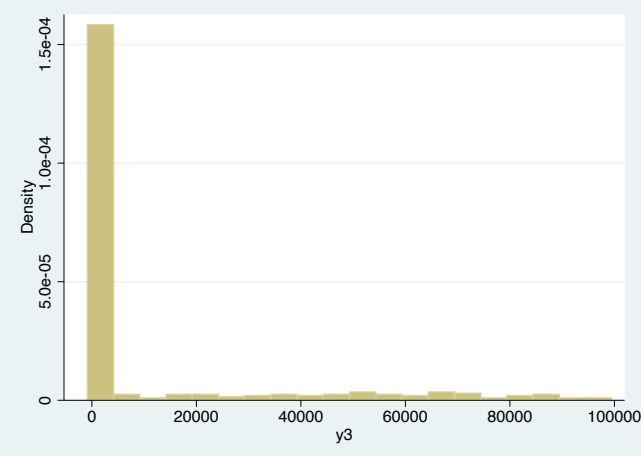
Limitation => Truncated data (only y)

. histogram y2



Limitation => Censored data

. histogram y3



Limitation => outlier

2. Estimate the model (1) for y_{1i}, y_{2i}, y_{3i} using OLS, using truncated regression model, Tobit model, determine the most appropriated models for each y_{ki}.

for y1

```
summarize y1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          y1 |        297    423.1683    227.8976   103.4663   1049.314

. reg y1 x

      Source |       SS           df       MS      Number of obs   =       297
-------------+----------------------------------   F(1, 295)       =     30.30
       Model |  1431905.59         1  1431905.59   Prob > F        =    0.0000
    Residual |  13941534.7       295  47259.4396   R-squared       =    0.0931
-------------+----------------------------------   Adj R-squared   =    0.0901
       Total |  15373440.3       296  51937.2982   Root MSE        =    217.39

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   123.1693   22.37637     5.50   0.000     79.13176    167.2069
       _cons |   69.01677   65.56422     1.05   0.293    -60.01611    198.0496
------------------------------------------------------------------------------

. predict y1_hat
(option xb assumed; fitted values)

. est store m_y1hat

. truncreg y1 x, ll(100) nolog
(note: 0 obs. truncated)

Truncated regression
Limit:   lower =        100                     Number of obs     =        297
         upper =       +inf                     Wald chi2(1)      =      23.70
Log likelihood =  -1975.456                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   243.9104   50.10515     4.87   0.000     145.7061    342.1147
       _cons |  -431.5919   172.0405    -2.51   0.012    -768.7851   -94.39859
-------------+----------------------------------------------------------------
      /sigma |   302.9879   25.65687    11.81   0.000     252.7013    353.2744
------------------------------------------------------------------------------

. predict truncated, e(0,.)

. est store m_y1trunc

. lrtest m_y1hat m_y1trunc, force

Likelihood-ratio test                                 LR chi2(1)  =     86.66
(Assumption: m_y1hat nested in m_y1trunc)             Prob > chi2 =    0.0000
```

from the limitation that we determined in question 1, we should use truncated regression model to solve the problem. And according to LR-test we can also reject the H0 which mean that the truncated regression model is the appropriated one.

For y2

```
. reg y2 x

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(1, 398)       =     76.58
       Model |  4567087.34         1  4567087.34   Prob > F        =    0.0000
    Residual |  23737174.2       398  59641.1413   R-squared       =    0.1614
-------------+----------------------------------   Adj R-squared   =    0.1592
       Total |  28304261.6       399  70937.9989   Root MSE        =    244.22

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   179.8325   20.55046     8.75   0.000     139.4315    220.2335
       _cons |  -179.3679   58.08821    -3.09   0.002     -293.566    -65.1698
------------------------------------------------------------------------------

. predict y2_hat
(option xb assumed; fitted values)

. est store m_y2hat

. tobit y2 x, ll(0)

Tobit regression                                  Number of obs     =        400
                                                  LR chi2(1)        =      73.17
                                                  Prob > chi2       =     0.0000
Log likelihood = -2389.3077                       Pseudo R2         =     0.0151

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   219.1562   24.86468     8.81   0.000     170.2741    268.0383
       _cons |  -318.6265   70.93087    -4.49   0.000    -458.0714   -179.1816
-------------+----------------------------------------------------------------
      /sigma |   284.9073   11.56717                      262.1671    307.6475
------------------------------------------------------------------------------
           72  left-censored observations at y2 <= 0
          328     uncensored observations
            0 right-censored observations

. predict censored, ystar(0,.)

. est store m_y2censor

. lrtest m_y2hat m_y2censor, force

Likelihood-ratio test                                 LR chi2(1)  =    752.97
(Assumption: m_y2hat nested in m_y2censor)            Prob > chi2 =    0.0000
```

from the limitation that we determined in question 1, we should use Tobit model to solve the problem. And according to LR-test we can also reject the H0 which mean that the Tobit model is the appropriated one.

.

For y3

```
reg y3 x

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(1, 398)       =     26.43
       Model |  1.3359e+10         1  1.3359e+10   Prob > F        =    0.0000
    Residual |  2.0115e+11       398   505402183   R-squared       =    0.0623
-------------+----------------------------------   Adj R-squared   =    0.0599
       Total |  2.1451e+11       399   537616802   Root MSE        =     22481

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   9726.042   1891.765     5.14   0.000     6006.941    13445.14
       _cons |  -16431.39   5347.288    -3.07   0.002    -26943.85   -5918.932
------------------------------------------------------------------------------
```

```
. predict y3_hat
(option xb assumed; fitted values)

. est store m_y3hat

. tobit y3 x, ul(2000)

Tobit regression                                  Number of obs     =        400
                                                  LR chi2(1)        =      52.03
                                                  Prob > chi2       =     0.0000
Log likelihood = -2693.1735                       Pseudo R2         =     0.0096

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   581.8005   79.00741     7.36   0.000     426.4777    737.1233
       _cons |  -961.6042   221.4967    -4.34   0.000    -1397.051   -526.1579
-------------+----------------------------------------------------------------
      /sigma |   911.3804   39.08073                      834.5505    988.2103
------------------------------------------------------------------------------
             0  left-censored observations
           315     uncensored observations
            85 right-censored observations at y3 >= 2000

. predict y3censored
(option xb assumed; fitted values)

. est store m_y3censor

. lrtest m_y3hat m_y3censor, force

Likelihood-ratio test                                 LR chi2(1)  =   3763.14
(Assumption: m_y3hat nested in m_y3censor)            Prob > chi2 =    0.0000
```

.

from the limitation that we determined in question 1, we should use Tobit model with upper limit to solve the problem. And according to LR-test we can also reject the H0 which mean that the Tobit model is the appropriated one.