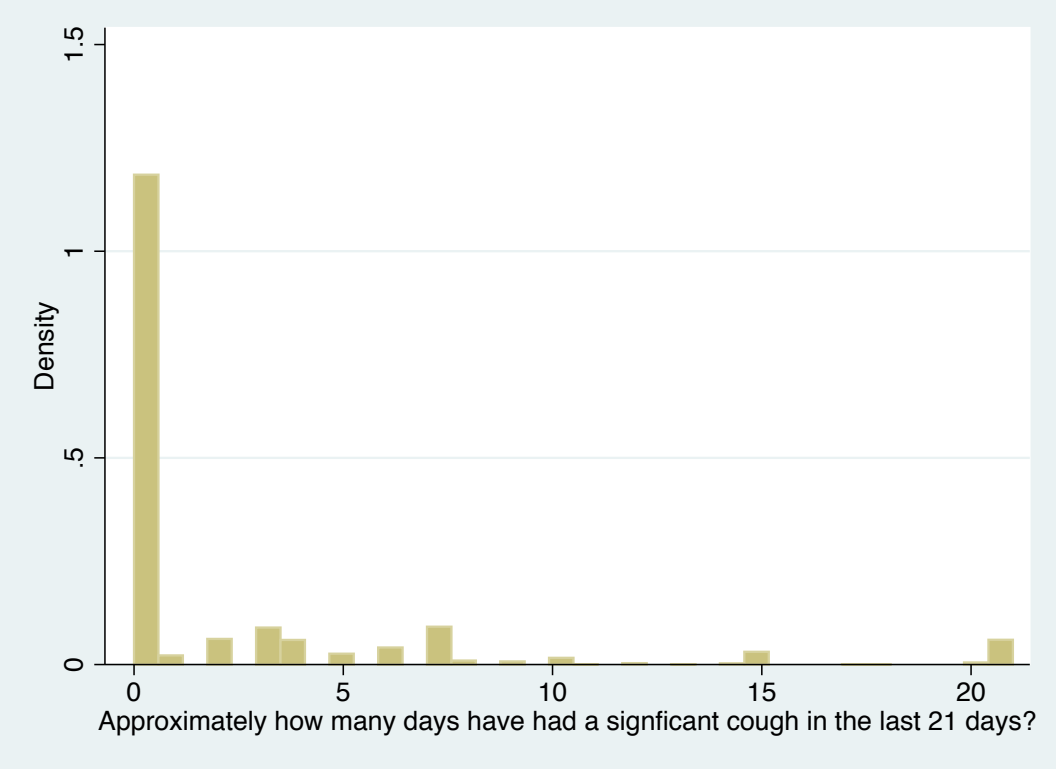
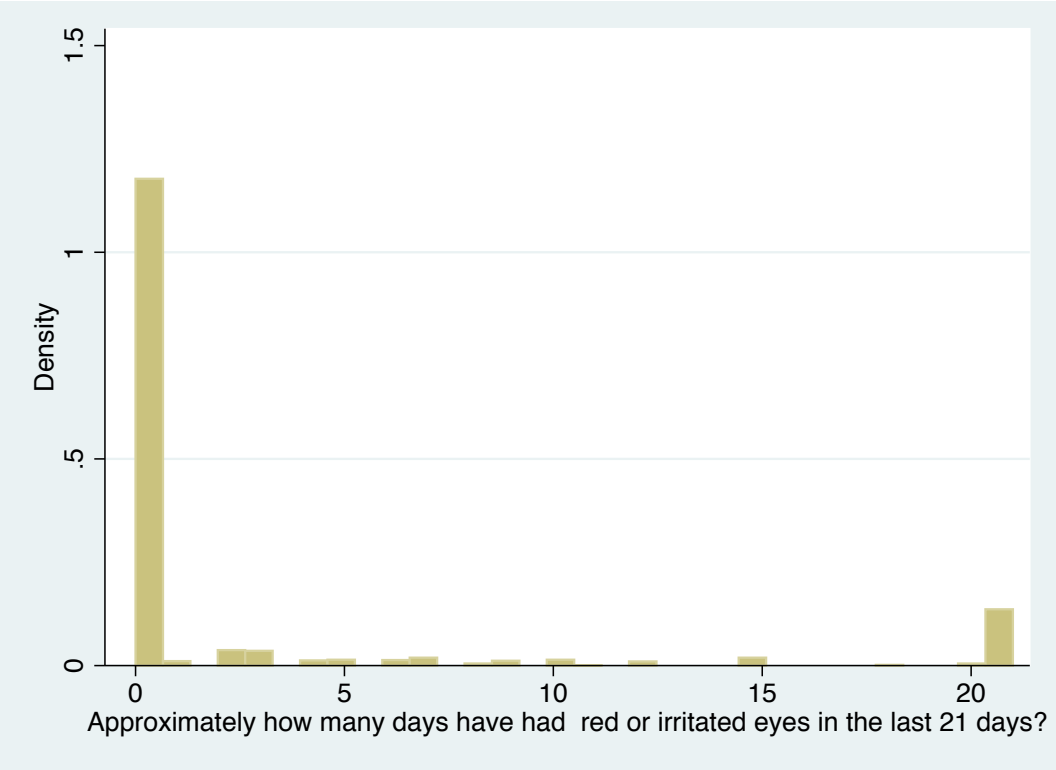
```
*1*
logit yi xi, noconstant
margins, dydx(*)
//The results show that coefficients are the same as the excel results.//
clear
*2*
use "/Users/aminatacissokho/Downloads/woodstove-2.dta"
//a//
hist coughdays
//Looking at the histogram, I observed that the distribution is skewed to the left. I see no evidence that the coughdays is also censored from above because there is less cough report from 20 to 21 days.//

*b*
tobit coughdays stove1 age age2 female year2009, ll(0)
tobit coughdays stove1 age age2 female year2009, ll(0) ul(21)
//The standard errors are high in the data censored from below and above compared to the data censored from only below.//
//These estimates differ because in the first regression, the data was only censored from below. However, in the second regression, the data was censored both below and above.//


*c*
hist redeyedays
//Looking at the above histogram showing the distribution of redeyedays, we can see the censoring in the data, that is, there are far more cases from 0 to 1 and 20.1 to 21.//
tobit redeyedays stove1 age age2 female year2009, ll(0)
tobit redeyedays stove1 age age2 female year2009, ll(0) ul(21)
//These estimates differ because in the first regression, the data was only censored from below. However, in the second regression, the data was censored both below and above.//

*3*
clear
use "/Users/aminatacissokho/Downloads/woodstove-2.dta"
//a//
regress coughdays stove1 age age2 female year2009 if lotteryhh==1
//b//
heckman coughdays stove1 age age2 female year2009, select ( lotterynumber)twostep
//The standard errors from the heckman regression are higher than that of the OLS. However, the coefficients varie in both regressions. The variable stove1 coefficient in the OLS regression is high but decreased in the heckman.//

//There is a difference in these results because the Ols regression doesn't differentiate between observable and unobservable data, which might lead to omitted variable bias, while the heckman regression uses a probit to identify whether a variable was selected or not and therefore adjust for selection bias.//

*c*
///No the lamda is not significant. The significance of the significance of the Lamda tells us  that there a selection bias issue in affecting our results.//

*d*
//Heckman two-step estimation breaks the estimation of the unconditional marginal effect into seperate problems (selection
```

```
unconditional marginal effect into seperate problems (selection
and behavioral), and each component estimates seperate
coeffitients with different regressors.//
//The maximum likelihood estimator simultaneously estimate the
selection and behavioral coefficients as well as the coefficient
on the inverse mills ratio.//

*4*
//a//
reg stove1 lotteryhh age age2 female year2009
predict fsresid,residuals
reg coughdays stove1 age age2 female year2009 fsresid
//No stove1 is not endogenuous because the predicted residual from
the original regression is not significant.//

//b//
reg coughdays stove1 age age2 female year2009
estimates store ols
ivreg coughdays (stove1 = lotteryhh)
estimates store iv
hausman iv ols
//Since Prob > 0.05, we fail to reject the null hypothesis that
the coefficients are the same, and conclude that there is no
significant difference in the coefficients due to bias in OLS due
to endogeneity.//
//Yes this result is the same as in the regression test. Stove1
was not endogenuous, so IV is not appropriate.//

//c//
//Taking all of thes estimation techniques into consideration, the
best estimation technique to look at the imppact of the woodstoves
on coughing is the heckman estimation because the coughing data
was censored from 0 to 21.//


*5*
foreach var in educ weaknessdays diarrheadays coughdays mucusdays
redeyedays backpaindays faintdays feverdays{
loneway hh_id `var' if year==2008
}
//The interclass correlation within the households is lowest among
the weaknessdays and backpaindays and highest among the the
faintdays.//

//b//
xtreg repcoughing educ female age age2
xtreg repcoughing educ female age age2,fe
//The coefficients and the standard errors with the fixed effects
regression are higher than the coefficients and standard errors
from the random effects regression.//

*c*
xtreg repcoughing educ female age age2,fe
estimates store fe
xtreg repcoughing educ female age age2
estimates store re
hausman fe re
// From the result, I failed to re jectthe null, therefore the
random effect is unbiased and appropriate for this estimation.//


*6*
reg yi xi
 predict resid, resid
```

```
ren resid es
gen es_squared= es^2
label var es_squared "squared residuals"
set matsize 800
mkmat xi, matrix (x)
mkmat yi, matrix (y)
mkmat es, matrix (e)
matrix list e
matrix betas=(invsym(x'*x))*(x'*y)
matrix list betas
matrix bun=invsym(x'*x)
matrix list bun
mkmat es_squared, matrix (E_squareds)
matrix list E_squareds

//The robust standard errors are same as we calculated in excel//
```


Density
0
.5
1
1.5
0
5
10
15
20
Approximately how many days have had a signficant cough in the last 21 days?


Density
0
.5
1
1.5
0
5
10
15
20
Approximately how many days have had red or irritated eyes in the last 21 days?

```
      name:  <unnamed>
       log:  /Users/aminatacissokho/Desktop/Econometrics_Assignment#2.smcl
  log type:  smcl
 opened on:   6 Nov 2016, 15:32:01

 1 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

 2 . logit yi xi, noconstant

  Iteration 0:   log likelihood = -3.4657359
  Iteration 1:   log likelihood = -2.9964031
  Iteration 2:   log likelihood = -2.9963636
  Iteration 3:   log likelihood = -2.9963636

  Logistic regression                             Number of obs     =          5
                                                  Wald chi2(1)      =       0.80
  Log likelihood = -2.9963636                     Prob > chi2       =     0.3717

------------------------------------------------------------------------------
          yi |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          xi |   .2835405   .3173817     0.89   0.372    -.3385162    .9055973
------------------------------------------------------------------------------

 3 .
  end of do-file

 4 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

 5 . margins, dydx(*)

  Average marginal effects                        Number of obs     =          5
  Model VCE    : OIM

  Expression   : Pr(yi), predict()
  dy/dx w.r.t. : xi

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          xi |   .0582299   .0422965     1.38   0.169    -.0246697    .1411294
------------------------------------------------------------------------------

 6 .
  end of do-file

 7 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

 8 . //The results show that coefficients are the same as the excel results.//
 9 .
  end of do-file

10 . clear

11 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

12 . use "/Users/aminatacissokho/Downloads/woodstove-2.dta"

13 .
  end of do-file

14 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

15 . //a//
```

```
16 . hist coughdays
   (bin=36, start=0, width=.58333333)

17 . //Looking at the histogram, I observed that the distribution is skewed to the
   >  left. I see no evidence that the coughdays is also censored from above becau
   > se there is less cough report from 20 to 21 days.//
18 .
   end of do-file

19 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

20 . *b*
21 . tobit coughdays stove1 age age2 female year2009, ll(0)

   Tobit regression                                  Number of obs     =      4,113
                                                     LR chi2(5)        =     219.83
                                                     Prob > chi2       =     0.0000
   Log likelihood = -6210.3645                       Pseudo R2         =     0.0174

   ------------------------------------------------------------------------------
      coughdays |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
   -------------+----------------------------------------------------------------
         stove1 |  -1.566984   .9408496    -1.67   0.096    -3.411559    .2775904
            age |  -.1633733   .0353096    -4.63   0.000    -.2325994   -.0941473
           age2 |   .0044113   .0005132     8.60   0.000     .0034052    .0054174
         female |   .9714853   .4238912     2.29   0.022     .1404289    1.802542
       year2009 |   1.807191   .4387403     4.12   0.000     .9470228     2.66736
          _cons |  -6.778135   .5818342   -11.65   0.000    -7.918845   -5.637425
   -------------+----------------------------------------------------------------
          /sigma |   10.97914   .2513457                      10.48637    11.47191
   ------------------------------------------------------------------------------
       2,844  left-censored observations at coughdays <= 0
       1,269     uncensored observations
           0 right-censored observations

22 . tobit coughdays stove1 age age2 female year2009, ll(0) ul(21)

   Tobit regression                                  Number of obs     =      4,113
                                                     LR chi2(5)        =     229.23
                                                     Prob > chi2       =     0.0000
   Log likelihood =  -5979.571                       Pseudo R2         =     0.0188

   ------------------------------------------------------------------------------
      coughdays |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
   -------------+----------------------------------------------------------------
         stove1 |  -1.742879    1.07017    -1.63   0.103    -3.840991    .3552332
            age |  -.1934421   .0404644    -4.78   0.000    -.2727742   -.1141101
           age2 |   .0052181   .0005929     8.80   0.000     .0040557    .0063805
         female |   1.093104   .4830594     2.26   0.024     .1460458    2.040162
       year2009 |   1.917816   .4998365     3.84   0.000     .9378657    2.897766
          _cons |  -7.814795    .668374   -11.69   0.000     -9.12517    -6.50442
   -------------+----------------------------------------------------------------
          /sigma |    12.4331   .3113713                      11.82264    13.04355
   ------------------------------------------------------------------------------
       2,844  left-censored observations at coughdays <= 0
       1,125     uncensored observations
         144 right-censored observations at coughdays >= 21

23 . //The standard errors are high in the data censored from below and above comp
   > ared to the data censored from only below.//
24 . //These estimates differ because in the first regression, the data was only c
   > ensored from below. However, in the second regression, the data was censored
   > both below and above.//
25 .
   end of do-file
```

```
26 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

27 . *c*
28 . hist redeyedays
   (bin=32, start=0, width=.65625)

29 . //Looking at the above histogram showing the distribution of redeyedays, we c
   > an see the censoring in the data, that is, there are far more cases from 0 to
   >  1 and 20.1 to 21.//
30 . tobit redeyedays stove1 age age2 female year2009, ll(0)

   Tobit regression                                  Number of obs     =      1,833
                                                     LR chi2(5)        =     163.27
                                                     Prob > chi2       =     0.0000
   Log likelihood = -2328.6878                       Pseudo R2         =     0.0339

------------------------------------------------------------------------------
  redeyedays |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      stove1 |  -4.504098   2.976655    -1.51   0.130    -10.3421    1.333904
         age |    .638936   .1719201     3.72   0.000     .3017555    .9761165
        age2 |  -.0023851   .0018411    -1.30   0.195    -.005996    .0012259
      female |   2.322365   1.135641     2.04   0.041     .0950752    4.549656
    year2009 |  -2.297421    1.16085    -1.98   0.048    -4.574153   -.0206885
       _cons |  -33.36741   3.924482    -8.50   0.000    -41.06435   -25.67047
-------------+----------------------------------------------------------------
       /sigma |   17.90802   .7341395                      16.46818    19.34786
------------------------------------------------------------------------------
        1,417  left-censored observations at redeyedays <= 0
          416     uncensored observations
            0 right-censored observations

31 . tobit redeyedays stove1 age age2 female year2009, ll(0) ul(21)

   Tobit regression                                  Number of obs     =      1,833
                                                     LR chi2(5)        =     165.45
                                                     Prob > chi2       =     0.0000
   Log likelihood = -1917.8266                       Pseudo R2         =     0.0414

------------------------------------------------------------------------------
  redeyedays |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      stove1 |  -7.698768    4.95948    -1.55   0.121    -17.42561    2.028075
         age |   1.001496   .2901903     3.45   0.001     .4323568    1.570636
        age2 |   -.003208   .0030984    -1.04   0.301    -.0092847    .0028688
      female |   3.872194   1.902353     2.04   0.042     .1411795    7.603208
    year2009 |  -3.961597    1.94314    -2.04   0.042    -7.772604   -.1505899
       _cons |  -54.85047   6.881848    -7.97   0.000    -68.34758   -41.35335
-------------+----------------------------------------------------------------
       /sigma |   29.24651   1.657097                      25.99651    32.49652
------------------------------------------------------------------------------
        1,417  left-censored observations at redeyedays <= 0
          252     uncensored observations
          164 right-censored observations at redeyedays >= 21

32 . //These estimates differ because in the first regression, the data was only c
   > ensored from below. However, in the second regression, the data was censored
   > both below and above.//
33 .
   end of do-file

34 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

35 . *3*
36 . clear
```

```
37 . use "/Users/aminatacissokho/Downloads/woodstove-2.dta"

38 . //a//
39 . regress coughdays stove1 age age2 female year2009 if lotteryhh==1
```

Source	SS	df	MS	Number of obs	=	1,006
				F(5, 1000)	=	18.14
Model	1917.7065	5	383.541299	Prob > F	=	0.0000
Residual	21145.4436	1,000	21.1454436	R-squared	=	0.0832
				Adj R-squared	=	0.0786
Total	23063.1501	1,005	22.9484081	Root MSE	=	4.5984

coughdays	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
stove1	-.8367116	.4441717	-1.88	0.060	-1.708327	.0349039
age	-.0734062	.0251452	-2.92	0.004	-.1227495	-.0240628
age2	.0021442	.0003886	5.52	0.000	.0013816	.0029068
female	.8659565	.2904885	2.98	0.003	.2959196	1.435993
year2009	-.1636158	.3398015	-0.48	0.630	-.8304215	.5031899
_cons	1.931123	.3581563	5.39	0.000	1.228299	2.633947

```
40 .
   end of do-file

41 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

42 . //b//
43 . heckman coughdays stove1 age age2 female year2009, select (lotterynumber)twos
   > tep
   note: two-step estimate of rho = -1.7722387 is being truncated to -1
```

Heckman selection model -- two-step estimates (regression model with sample selection)

Number of obs	=	1,052
Censored obs	=	25
Uncensored obs	=	1,027
Wald chi2(5)	=	24.05
Prob > chi2	=	0.0002

coughdays	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
coughdays						
stove1	-1.167674	.9777138	-1.19	0.232	-3.083958	.7486097
age	-.0704107	.0515089	-1.37	0.172	-.1713663	.0305449
age2	.0021707	.0007951	2.73	0.006	.0006123	.0037292
female	.8597896	.5940217	1.45	0.148	-.3044715	2.024051
year2009	-.0454716	.6935064	-0.07	0.948	-1.404719	1.313776
_cons	2.419931	.9468161	2.56	0.011	.5642054	4.275656
select						
lotterynumber	-.0106082	.003596	-2.95	0.003	-.0176562	-.0035602
_cons	2.544843	.2249026	11.32	0.000	2.104042	2.985644
mills						
lambda	-10.06733	10.79063	-0.93	0.351	-31.21657	11.0819
rho	-1.00000					
sigma	10.067334					

```
44 . //The standard errors from the heckman regression are higher than that of the
   >  OLS. However, the coefficients varie in both regressions. The variable stove
   > 1 coefficient in the OLS regression is high but decreased in the heckman.//
45 .
```

```
46 .
   end of do-file

47 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

48 . *c*
49 . ///No the lamda is not significant. The significance of the significance of t
   > he Lamda tells us  that there a selection bias issue in affecting our results
   > .//

   end of do-file

50 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

51 . *d*
52 . //Heckman two-step estimation breaks the estimation of the unconditional marg
   > inal effect into seperate problems (selection and behavioral), and each compo
   > nent estimates seperate coeffitients with different regressors.//
53 . //The maximum likelihood estimator simultaneously estimate the selection and
   > behavioral coefficients as well as the coefficient on the inverse mills ratio
   > .//
54 .
   end of do-file

55 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

56 . *4*
57 . //a//
58 . reg stove1 lotteryhh age age2 female year2009
```

Source	SS	df	MS	Number of obs	=	4,172
				F(5, 4166)	=	155.83
Model	35.9171943	5	7.18343885	Prob > F	=	0.0000
Residual	192.045414	4,166	.046098275	R-squared	=	0.1576
				Adj R-squared	=	0.1565
Total	227.962608	4,171	.054654186	Root MSE	=	.21471

stove1	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lotteryhh	.1632292	.0077206	21.14	0.000	.1480927	.1783657
age	-.0009034	.0005748	-1.57	0.116	-.0020303	.0002234
age2	9.38e-06	8.60e-06	1.09	0.275	-7.47e-06	.0000262
female	.0010354	.0066512	0.16	0.876	-.0120044	.0140752
year2009	.1236397	.0066557	18.58	0.000	.1105911	.1366884
_cons	-.031034	.0087233	-3.56	0.000	-.0481364	-.0139316

```
59 . predict fsresid,residuals
   (13 missing values generated)

60 . reg coughdays stove1 age age2 female year2009 fsresid
```

Source	SS	df	MS	Number of obs	=	4,113
				F(6, 4106)	=	60.03
Model	7642.15165	6	1273.69194	Prob > F	=	0.0000
Residual	87120.0662	4,106	21.2177463	R-squared	=	0.0806
				Adj R-squared	=	0.0793
Total	94762.2178	4,112	23.0452864	Root MSE	=	4.6063

coughdays	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
stove1	-.300265	1.025303	-0.29	0.770	-2.310414	1.709884
age	-.062733	.0124676	-5.03	0.000	-.0871763	-.0382898
age2	.0019652	.0001857	10.58	0.000	.0016011	.0023293

```
      female |   .3251216   .1437175     2.26   0.024     .0433575    .6068857
    year2009 |   .1095251   .1888229     0.58   0.562    -.2606701    .4797203
     fsresid |  -.5112316    1.07802    -0.47   0.635    -2.624734    1.602271
       _cons |   1.901167   .1825497    10.41   0.000      1.54327    2.259063
------------------------------------------------------------------------------

61 . //No stove1 is not endogenuous because the predicted residual from the origin
   > al regression is not significant.//
62 .
   end of do-file

63 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

64 . //b//
65 . reg coughdays stove1 age age2 female year2009

      Source |       SS           df       MS      Number of obs   =     4,113
-------------+----------------------------------   F(5, 4107)      =     72.00
       Model |  7637.37986         5  1527.47597   Prob > F        =    0.0000
    Residual |   87124.838     4,107  21.2137419   R-squared       =    0.0806
-------------+----------------------------------   Adj R-squared   =    0.0795
       Total |  94762.2178     4,112  23.0452864   Root MSE        =    4.6058

------------------------------------------------------------------------------
   coughdays |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      stove1 |  -.7621269   .3204928    -2.38   0.017    -1.390466   -.1337874
         age |  -.0633158   .0124057    -5.10   0.000    -.0876377   -.0389939
        age2 |   .0019709   .0001853    10.64   0.000     .0016076    .0023342
      female |    .325472    .143702     2.26   0.024     .0437383    .6072058
    year2009 |   .1645739    .148914     1.11   0.269    -.1273781    .4565259
       _cons |   1.908721   .1818362    10.50   0.000     1.552223    2.265218
------------------------------------------------------------------------------

66 . estimates store ols

67 . ivreg coughdays (stove1 = lotteryhh)

Instrumental variables (2SLS) regression

      Source |       SS           df       MS      Number of obs   =     4,124
-------------+----------------------------------   F(1, 4122)      =      0.62
       Model |  126.562678         1  126.562678   Prob > F        =    0.4299
    Residual |  94862.0261     4,122   23.013592   R-squared       =    0.0013
-------------+----------------------------------   Adj R-squared   =    0.0011
       Total |  94988.5887     4,123  23.0387069   Root MSE        =    4.7972

------------------------------------------------------------------------------
   coughdays |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      stove1 |  -.8728325   1.105524    -0.79   0.430    -3.040255     1.29459
       _cons |   2.331502   .0977189    23.86   0.000     2.139921    2.523084
------------------------------------------------------------------------------
Instrumented:  stove1
Instruments:   lotteryhh
------------------------------------------------------------------------------

68 . estimates store iv

69 . hausman iv ols

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       iv          ols         Difference          S.E.
-------------+----------------------------------------------------------------
      stove1 |  -.8728325   -.7621269       -.1107056        1.058048
```

```
-------------+----------------------------------------------------------------
                            b = consistent under Ho and Ha; obtained from ivreg
             B = inconsistent under Ha, efficient under Ho; obtained from regress

    Test:  Ho:  difference in coefficients not systematic

                  chi2(1) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        0.01
                Prob>chi2 =      0.9167

70 . //Since Prob > 0.05, we fail to reject the null hypothesis that the coefficie
   > nts are the same, and conclude that there is no significant difference in the
   >  coefficients due to bias in OLS due to endogeneity.//
71 . //Yes this result is the same as in the regression test. Stove1 was not endog
   > enuous, so IV is not appropriate.//
72 .
   end of do-file

73 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

74 . //c//
75 . //Taking all of thes estimation techniques into consideration, the best estim
   > ation technique to look at the imppact of the woodstoves on coughing is the h
   > eckman estimation because the coughing data was censored from 0 to 21.//
76 .
   end of do-file

77 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

78 . *5*
79 . foreach var in educ weaknessdays diarrheadays coughdays mucusdays redeyedays
   > backpaindays faintdays feverdays{
     2. loneway hh_id `var' if year==2008
     3. }

            One-way Analysis of Variance for hh_id: household id number

                                                Number of obs =        2,145
                                                    R-squared =       0.0096

    Source                SS         df      MS            F     Prob > F
-----------------------------------------------------------------------------
Between educ           919714.84      16     57482.178      1.29     0.1925
Within educ             94649830   2,128     44478.304
-----------------------------------------------------------------------------
Total                   95569545   2,144     44575.347

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ---------------------------------------------------
            0.00257     0.00416        0.00000     0.01073

         Estimated SD of educ effect              10.70966
         Estimated SD within educ                 210.8988
         Est. reliability of a educ mean           0.22622
              (evaluated at n=113.38)

            One-way Analysis of Variance for hh_id: household id number

                                                Number of obs =        2,148
                                                    R-squared =       0.0074

    Source                SS         df      MS            F     Prob > F
-----------------------------------------------------------------------------
Between weaknessdays   704862.97      18     39159.054      0.88     0.6060
Within weaknessdays     94938853   2,129     44593.167
```

```
------------------------------------------------------------------------------
Total                   95643716  2,147    44547.609

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ---------------------------------------------------
            0.00000*     0.00575        0.00000     0.01127

         Estimated SD of weaknessdays effect            .
         Estimated SD within weaknessdays         211.1709
         Est. reliability of a weaknessdays mean   0.00000*
              (evaluated at n=58.21)

(*) Truncated at zero.

          One-way Analysis of Variance for hh_id: household id number

                                              Number of obs =        2,148
                                                  R-squared =       0.0113

    Source                SS         df      MS            F     Prob > F
------------------------------------------------------------------------------
Between diarrheadays    1078857.5     15     71923.835      1.62     0.0608
Within diarrheadays      94564859  2,132     44354.999
------------------------------------------------------------------------------
Total                   95643716  2,147     44547.609

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ---------------------------------------------------
            0.01113     0.01272        0.00000     0.03607

         Estimated SD of diarrheadays effect      22.34266
         Estimated SD within diarrheadays         210.6063
         Est. reliability of a diarrheadays mean   0.38331
              (evaluated at n=55.23)

          One-way Analysis of Variance for hh_id: household id number

                                              Number of obs =        2,148
                                                  R-squared =       0.0132

    Source                SS         df      MS            F     Prob > F
------------------------------------------------------------------------------
Between coughdays       1262034.2     18     70113.009      1.58     0.0564
Within coughdays         94381682  2,129     44331.462
------------------------------------------------------------------------------
Total                   95643716  2,147     44547.609

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ---------------------------------------------------
            0.01062     0.01170        0.00000     0.03355

         Estimated SD of coughdays effect         21.81554
         Estimated SD within coughdays            210.5504
         Est. reliability of a coughdays mean      0.36771
              (evaluated at n=54.17)

          One-way Analysis of Variance for hh_id: household id number

                                              Number of obs =        2,148
                                                  R-squared =       0.0154

    Source                SS         df      MS            F     Prob > F
------------------------------------------------------------------------------
```

Between mucusdays	1476450.4	15	98430.024	2.23	0.0043
Within mucusdays	94167266	2,132	44168.511		
Total	95643716	2,147	44547.609		

Intraclass correlation	Asy. S.E.	[95% Conf.	Interval]
0.02336	0.02133	0.00000	0.06518

Estimated SD of mucusdays effect 32.50622
Estimated SD within mucusdays 210.1631
Est. reliability of a mucusdays mean 0.55127
(evaluated at n=51.35)

One-way Analysis of Variance for hh_id: household id number

Number of obs = 938
R-squared = 0.0240

Source	SS	df	MS	F	Prob > F
Between redeyedays	1001804.4	15	66786.961	1.51	0.0943
Within redeyedays	40761471	922	44209.838		
Total	41763275	937	44571.265		

Intraclass correlation	Asy. S.E.	[95% Conf.	Interval]
0.01976	0.02588	0.00000	0.07048

Estimated SD of redeyedays effect 29.85125
Estimated SD within redeyedays 210.2614
Est. reliability of a redeyedays mean 0.33805
(evaluated at n=25.34)

One-way Analysis of Variance for hh_id: household id number

Number of obs = 941
R-squared = 0.0141

Source	SS	df	MS	F	Prob > F
Between backpaindays	589294.92	17	34664.407	0.78	0.7212
Within backpaindays	41174625	923	44609.562		
Total	41763920	940	44429.702		

Intraclass correlation	Asy. S.E.	[95% Conf.	Interval]
0.00000*	0.01058	0.00000	0.02074

Estimated SD of backpaindays effect .
Estimated SD within backpaindays 211.2098
Est. reliability of a backpaindays mean 0.00000*
(evaluated at n=32.72)

(*) Truncated at zero.

One-way Analysis of Variance for hh_id: household id number

Number of obs = 941
R-squared = 0.0313

```
    Source                SS         df      MS            F     Prob > F
-----------------------------------------------------------------------------
Between faintdays        1306765      16     81672.815      1.87     0.0202
Within faintdays        40457155     924      43784.8
-----------------------------------------------------------------------------
Total                   41763920     940     44429.702

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ------------------------------------------------
            0.03844     0.03857        0.00000     0.11404

         Estimated SD of faintdays effect         41.83896
         Estimated SD within faintdays            209.2482
         Est. reliability of a faintdays mean      0.46390
              (evaluated at n=21.64)

          One-way Analysis of Variance for hh_id: household id number

                                              Number of obs =        2,144
                                                  R-squared =       0.0148

    Source                SS         df      MS            F     Prob > F
-----------------------------------------------------------------------------
Between feverdays       1418642.7      17     83449.573      1.88     0.0154
Within feverdays        94143205    2,126     44281.846
-----------------------------------------------------------------------------
Total                   95561848    2,143     44592.556

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ------------------------------------------------
            0.01682     0.01643        0.00000     0.04902

         Estimated SD of feverdays effect         27.52579
         Estimated SD within feverdays            210.4325
         Est. reliability of a feverdays mean      0.46936
              (evaluated at n=51.70)

80 . //The interclass correlation within the households is lowest among the weakne
   > ssdays and backpaindays and highest among the the faintdays.//
81 .
   end of do-file

82 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

83 . //b//
84 . xtreg repcoughing educ female age age2

   Random-effects GLS regression                   Number of obs     =      3,045
   Group variable: hh_id                           Number of groups  =        445

   R-sq:                                           Obs per group:
        within  = 0.0374                                         min =          1
        between = 0.0568                                         avg =        6.8
        overall = 0.0460                                         max =         18

                                                   Wald chi2(4)      =     127.50
   corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
 repcoughing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |  -.0092811   .0029558    -3.14   0.002    -.0150743   -.0034879
      female |     .03795   .0147473     2.57   0.010     .0090458    .0668541
         age |  -.0010465   .0013498    -0.78   0.438    -.0036922    .0015991
```

```
------------------------------------------------------------------------------
        age2 |   .0000753   .0000202     3.73   0.000     .0000357    .0001149
       _cons |   .2713831    .023255    11.67   0.000     .2258041     .316962
-------------+----------------------------------------------------------------
     sigma_u |  .23511343
     sigma_e |  .39247647
         rho |  .26409017   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

```
85 . xtreg repcoughing educ female age age2,fe

Fixed-effects (within) regression               Number of obs     =      3,045
Group variable: hh_id                           Number of groups  =        445

R-sq:                                           Obs per group:
     within  = 0.0375                                         min =          1
     between = 0.0550                                         avg =        6.8
     overall = 0.0454                                         max =         18

                                                F(4,2596)         =      25.30
corr(u_i, Xb)  = 0.0233                         Prob > F          =     0.0000

------------------------------------------------------------------------------
 repcoughing |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |  -.0085506   .0031982    -2.67   0.008    -.0148218   -.0022793
      female |   .0403892   .0150041     2.69   0.007     .010968    .0698105
         age |  -.0018897   .0014223    -1.33   0.184    -.0046787    .0008994
        age2 |   .0000861    .000022     3.92   0.000      .000043    .0001292
       _cons |   .2648548   .0205654    12.88   0.000     .2245285    .3051811
-------------+----------------------------------------------------------------
     sigma_u |  .29814141
     sigma_e |  .39247647
         rho |  .36590678   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(444, 2596) = 3.04                     Prob > F = 0.0000

86 . //The coefficients and the standard errors with the fixed effects regression
> are higher than the coefficients and standard errors from the random effects
> regression.//
87 .
end of do-file

88 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

89 . *c*
90 . xtreg repcoughing educ female age age2,fe

Fixed-effects (within) regression               Number of obs     =      3,045
Group variable: hh_id                           Number of groups  =        445

R-sq:                                           Obs per group:
     within  = 0.0375                                         min =          1
     between = 0.0550                                         avg =        6.8
     overall = 0.0454                                         max =         18

                                                F(4,2596)         =      25.30
corr(u_i, Xb)  = 0.0233                         Prob > F          =     0.0000

------------------------------------------------------------------------------
 repcoughing |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |  -.0085506   .0031982    -2.67   0.008    -.0148218   -.0022793
      female |   .0403892   .0150041     2.69   0.007     .010968    .0698105
         age |  -.0018897   .0014223    -1.33   0.184    -.0046787    .0008994
        age2 |   .0000861    .000022     3.92   0.000      .000043    .0001292
       _cons |   .2648548   .0205654    12.88   0.000     .2245285    .3051811
------------------------------------------------------------------------------
```

```
-------------+----------------------------------------------------------------
     sigma_u |  .29814141
     sigma_e |  .39247647
         rho |  .36590678   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(444, 2596) = 3.04                     Prob > F = 0.0000

91 . estimates store fe

92 . xtreg repcoughing educ female age age2

Random-effects GLS regression                   Number of obs     =       3,045
Group variable: hh_id                           Number of groups  =         445

R-sq:                                           Obs per group:
     within  = 0.0374                                         min =           1
     between = 0.0568                                         avg =         6.8
     overall = 0.0460                                         max =          18

                                                Wald chi2(4)      =      127.50
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =      0.0000

------------------------------------------------------------------------------
 repcoughing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |  -.0092811   .0029558    -3.14   0.002    -.0150743   -.0034879
      female |     .03795   .0147473     2.57   0.010     .0090458    .0668541
         age |  -.0010465   .0013498    -0.78   0.438    -.0036922    .0015991
        age2 |   .0000753   .0000202     3.73   0.000     .0000357    .0001149
       _cons |   .2713831    .023255    11.67   0.000     .2258041     .316962
-------------+----------------------------------------------------------------
     sigma_u |  .23511343
     sigma_e |  .39247647
         rho |  .26409017   (fraction of variance due to u_i)
------------------------------------------------------------------------------

93 . estimates store re

94 . hausman fe re

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       fe           re         Difference          S.E.
-------------+----------------------------------------------------------------
        educ |  -.0085506   -.0092811        .0007305        .0012214
      female |   .0403892      .03795        .0024393        .0027643
         age |  -.0018897   -.0010465       -.0008431        .0004483
        age2 |   .0000861    .0000753        .0000108        8.70e-06
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(4) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        6.51
                Prob>chi2 =      0.1639

95 . // From the result, I failed to re jectthe null, therefore the random effect
   > is unbiased and appropriate for this estimation.//
96 .
   end of do-file

97 . log close
       name:  <unnamed>
        log:  /Users/aminatacissokho/Desktop/Econometrics_Assignment#2.smcl
```

```
  log type:  smcl
closed on:   6 Nov 2016, 15:37:40
```

Logistic Regression

Xi	Yi
1	0
2	1
3	0
4	1
5	1

X-bar

3

Pi

0.700697511

(1-Pi)

0.299302489

Xi	*Success*	*Failure*	*Total*	*p-Pred*	*LL*
	0	1	1	0.570414	-0.84493
	1	0	1	0.638089	-0.44928
	0	1	1	0.700698	-1.2063
	1	0	1	0.756606	-0.27891
	1	0	1	0.804978	-0.21694
Total	3	2	5		-2.99636

Coeff

0.283541 B1

Marginal effects
0.059464

```
      name:  <unnamed>
       log:  /Users/aminatacissokho/Desktop/Qestion6.smcl
  log type:  smcl
 opened on:   6 Nov 2016, 16:13:52

1 . do "/var/folders/87/_v_xwmn95134jf_7wn731xqr0000gn/T//SD00816.000000"

2 . reg yi xi

      Source |       SS           df       MS      Number of obs   =         5
-------------+----------------------------------   F(1, 3)         =      1.50
       Model |          .4         1          .4   Prob > F        =    0.3081
    Residual |          .8         3  .266666667   R-squared       =    0.3333
-------------+----------------------------------   Adj R-squared   =    0.1111
       Total |         1.2         4          .3   Root MSE        =     .5164

------------------------------------------------------------------------------
          yi |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          xi |          .2   .1632993     1.22   0.308    -.3196913    .7196913
       _cons |  -2.22e-16   .5416026    -0.00   1.000    -1.723621    1.723621
------------------------------------------------------------------------------

3 .  predict resid, resid

4 .  ren resid es
variable es already defined
r(110);

end of do-file

r(110);

5 . clear

6 . *(2 variables, 5 observations pasted into data editor)

7 . log close
      name:  <unnamed>
       log:  /Users/aminatacissokho/Desktop/Qestion6.smcl
  log type:  smcl
 closed on:   6 Nov 2016, 16:15:23
```