```
  ___  ____  ____  ____  ____ (R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/   14.0   Copyright 1985-2015 StataCorp LP
  Statistics/Data Analysis            StataCorp
                                      4905 Lakeway Drive
     MP - Parallel Edition            College Station, Texas 77845 USA
                                      800-STATA-PC        http://www.stata.com
                                      979-696-4600        stata@stata.com
                                      979-696-4601 (fax)

Single-user 8-core Stata perpetual license:
       Serial number:  10699393
         Licensed to:  Andong Yan
                       University of Southern California

Notes:
      1.  Unicode is supported; see help unicode advice.
      2.  More than 2 billion observations are allowed; see help obs advice.
      3.  Maximum number of variables is set to 5000; see help set maxvar.

. doedit "C:\Users\yadto\Dropbox\USC\ECON513\HW4\hw4 ye.do"

. do "C:\Users\yadto\Dropbox\USC\ECON513\HW4\hw4 ye.do"

. use "C:\Users\yadto\Dropbox\USC\ECON513\HW4\Card-Krueger.dta", clear

.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
```

```
.
.
. //--------------------question a ----------------------
. // generate delta y
. gen deltay = EMPL1 -EMPL0

.
. //generate clusters by chains
. gen bk = 0

. replace bk = 1 if CHAIN == 1
(163 real changes made)

. gen kfc = 0

. replace kfc = 1 if CHAIN == 2
(80 real changes made)

. gen roy = 0

. replace roy = 1 if CHAIN == 3
(95 real changes made)

. gen wend = 0

. replace wend = 1 if CHAIN ==4
(53 real changes made)

.
. // reg changeiny d bk kfc roy wend no intercept
. reg deltay STATE bk kfc roy wend, nocon

      Source |       SS           df       MS      Number of obs   =       391
-------------+----------------------------------   F(5, 386)       =      2.50
       Model |  959.586342         5  191.917268   Prob > F        =    0.0301
    Residual |  29596.2262       386   76.674161   R-squared       =    0.0314
-------------+----------------------------------   Adj R-squared   =    0.0189
       Total |  30555.8125       391  78.1478581   Root MSE        =    8.7564

------------------------------------------------------------------------------
      deltay |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       STATE |   2.965651    1.12227     2.64   0.009     .7591239    5.172177
          bk |  -2.051036   1.122164    -1.83   0.068    -4.257355    .1552819
         kfc |  -1.561428   1.366897    -1.14   0.254    -4.248924    1.126068
         roy |  -4.313903   1.286916    -3.35   0.001    -6.844145    -1.78366
        wend |  -2.110868   1.471083    -1.43   0.152    -5.003206    .7814698
------------------------------------------------------------------------------

. // coefficient of STATE is what we are interested
. // compared to result in lecture 11, p9, beta is slightly bigger in condiftional
. // --> dif-in-dif, and standard error seems to be no change.
.
```

```
.
. //----------------------question b----------------------
. //store the residuals
. predict e, r

.
. //construct matrix Z, indicator dtc, t time, c chain
. gen d01 = 0

. gen d02 = 0

. gen d03 = 0

. gen d04 = 0

. gen d11 = 0

. gen d12 = 0

. gen d13 = 0

. gen d14 = 0

. replace d01 = 1 if STATE == 0 & bk == 1
(34 real changes made)

. replace d02 = 1 if STATE == 0 & kfc == 1
(12 real changes made)

. replace d03 = 1 if STATE == 0 & roy == 1
(17 real changes made)

. replace d04 = 1 if STATE == 0 & wend == 1
(13 real changes made)

. replace d11 = 1 if STATE == 1 & bk == 1
(129 real changes made)

. replace d12 = 1 if STATE == 1 & kfc == 1
(68 real changes made)

. replace d13 = 1 if STATE == 1 & roy == 1
(78 real changes made)

. replace d14 = 1 if STATE == 1 & wend == 1
(40 real changes made)

. mkmat d01 d02 d03 d04 d11 d12 d13 d14, mat(z) //this is matirx z

. mkmat e, mat(e) // this is vector e

.
. //number of observations in each cluster
```

```
. forval i =1/4 {
  2. qui sum d0`i'
  3. qui gen n0`i' = r(sum)
  4. qui sum d1`i'
  5. qui gen n1`i' = r(sum)
  6. }

. sum n01 n02 n03 n04 n11 n12 n13 n14

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         n01 |       391          34           0         34         34
         n02 |       391          12           0         12         12
         n03 |       391          17           0         17         17
         n04 |       391          13           0         13         13
         n11 |       391         129           0        129        129
-------------+--------------------------------------------------------
         n12 |       391          68           0         68         68
         n13 |       391          78           0         78         78
         n14 |       391          40           0         40         40

.
.
. mata:
------------------------------------------------- mata (type end to exit) ----------
: z = st_matrix("z")

: e = st_matrix("e")

: z2inv = invsym(z'*z)

: I = I(391)

: etilde = (I-z*z2inv*z')*e

: variance = (etilde'*etilde)/(391-8-1)

: st_matrix("var", variance)

: end
------------------------------------------------------------------------------------

.
. scalar vartilde = var[1,1]

.
. mat list var

symmetric var[1,1]
           c1
r1  76.781702

. sum e

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           e |       391    1.51e-08    8.711358  -41.44896   34.14522
```

```
. scalar sdbar = r(sd)

. scalar varbar = sdbar^2

. display varbar
75.887759

.
. //it appears that varbar is smaller than vartilde, which is weird, I take the
. // --> absolute value of their difference
.
. scalar rhohat = (vartilde - varbar)/varbar

. display rhohat
.01177981

. scalar varcluster = rhohat * varbar

. scalar varindio = (1-rhohat)*varbar

.
. // correction factor
. scalar L = 391/8

. scalar correct = L*rhohat + (1-rhohat)

. display correct
1.5639582

.
. //corrected standard error
. scalar correctvar = correct*rhohat

. scalar correctsd = sqrt(correctvar)

. display correctsd
.1357318

.
. //-------------------question d-------------------
.
.
. //I dont know how to set up matrix V, so I did not do this one
.
.
.
. //-------------------question e -------------------
. gen sbk = STATE*bk

. gen skfc = STATE*kfc

. gen sroy = STATE*roy
```

```
. gen swend = STATE*wend

. reg deltay STATE bk kfc roy wend sbk skfc sroy swend, nocon
note: swend omitted because of collinearity

      Source |       SS           df       MS      Number of obs   =       391
-------------+----------------------------------   F(8, 383)       =      2.00
       Model |  1225.20219         8  153.150274   Prob > F        =    0.0454
    Residual |  29330.6103       383   76.581228   R-squared       =    0.0401
-------------+----------------------------------   Adj R-squared   =    0.0200
       Total |  30555.8125       391  78.1478581   Root MSE        =    8.7511

------------------------------------------------------------------------------
      deltay |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       STATE |   3.583173   2.793814     1.28   0.200    -1.909959    9.076306
          bk |  -3.367647   1.500796    -2.24   0.025    -6.318478   -.4168159
         kfc |   2.041667   2.526216     0.81   0.419    -2.925322    7.008656
         roy |  -3.867647   2.122446    -1.82   0.069    -8.040752    .3054584
        wend |  -2.576923    2.42711    -1.06   0.289    -7.349052    2.195206
         sbk |   1.046102   3.263654     0.32   0.749     -5.37082    7.463024
        skfc |  -4.856457   3.913229    -1.24   0.215    -12.55056    2.837643
        sroy |  -1.161039   3.645817    -0.32   0.750    -8.329361    6.007284
       swend |          0  (omitted)
------------------------------------------------------------------------------

. //Yes, we can estimate the clustered standard errors since each variable is
. //--> cluster specific, so we still can calculate
.
.
. //------------------question f --------------------
.
. reg deltay STATE bk kfc roy wend, nocon vce(boot)
(running regress on estimation sample)

Bootstrap replications (50)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50

Linear regression                               Number of obs     =        391
                                                Replications      =         50
                                                Wald chi2(5)      =      15.10
                                                Prob > chi2       =     0.0099
                                                R-squared         =     0.0314
                                                Adj R-squared     =     0.0189
                                                Root MSE          =     8.7564

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
      deltay |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       STATE |   2.965651   1.511061     1.96   0.050     .0040244    5.927277
          bk |  -2.051036   1.598939    -1.28   0.200    -5.184899    1.082827
         kfc |  -1.561428   1.261104    -1.24   0.216    -4.033146    .9102901
         roy |  -4.313903   1.446553    -2.98   0.003    -7.149094   -1.478711
        wend |  -2.110868   1.744677    -1.21   0.226    -5.530372    1.308635
------------------------------------------------------------------------------

. // the standard error grows bigger compared to a
.
end of do-file

.
```