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 .

```
1 . doedit "D:\THU\Sophomore\Econometrics\PS4\ps4-cr.do"

2 . do "D:\THU\Sophomore\Econometrics\PS4\ps4-cr.do"

3 . // estimating model: we collect data. (y,x) => estimate beta
4 . // numeric simulation: we would set some beta. generate data (y,x| beta). estimate, compare with the
5 . cd "D:\THU\Sophomore\Econometrics\PS4"
D:\THU\Sophomore\Econometrics\PS4

6 . clear

7 .
8 . log using "2022011545-5.log", replace
-------------------------------------------------------------------------------
      name:  <unnamed>
       log:  D:\THU\Sophomore\Econometrics\PS4\2022011545-5.log
  log type:  text
 opened on:  21 May 2024, 22:02:40

9 .
10 . clear

11 . // Set number of observations
12 . eststo clear

13 . set obs 1000
number of observations (_N) was 0, now 1,000

14 . // Generate variables
15 . /// y = beta0 + beta1 * x + u.
> g id = _n

16 . g x = runiform()

17 . g u = rnormal()

18 . g y = 3*x + u

19 . // Estimate the model using OLS
20 . reg y x

      Source |       SS           df       MS      Number of obs   =         1,000
-------------+----------------------------------   F(1, 998)       =       808.43
       Model |  805.20678         1   805.20678   Prob > F        =       0.0000
    Residual |  994.01803       998   .99601005   R-squared       =       0.4475
-------------+----------------------------------   Adj R-squared   =       0.4470
       Total |  1799.22481       999  1.80102584   Root MSE        =         .998
```

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x	**3.082817**	**.1084241**	**28.43**	**0.000**	**2.870052**	**3.295583**
_cons	**-.0447124**	**.0622607**	**-0.72**	**0.473**	**-.1668892**	**.0774644**

```
21 . eststo m1 //store the results

22 .
23 . // Replicate each observation by 4
24 . expand 4, gen(dupind)
   (3,000 observations created)

25 .
26 . // Estimate the model using OLS and the new sample
27 . reg y x
```

Source	SS	df	MS	Number of obs	=	**4,000**
				F(1, 3998)	=	**3238.59**
Model	**3220.82712**	**1**	**3220.82712**	Prob > F	=	**0.0000**
Residual	**3976.07212**	**3,998**	**.994515288**	R-squared	=	**0.4475**
				Adj R-squared	=	**0.4474**
Total	**7196.89924**	**3,999**	**1.79967473**	Root MSE	=	**.99725**

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x	**3.082817**	**.0541714**	**56.91**	**0.000**	**2.976611**	**3.189023**
_cons	**-.0447124**	**.031107**	**-1.44**	**0.151**	**-.1056994**	**.0162746**

```
28 . eststo m2

29 . // Cluster the standard error
30 . reg y x, cluster(id)
```

Linear regression

Number of obs	=	**4,000**	
F(1, 999)	=	**851.18**	
Prob > F	=	**0.0000**	
R-squared	=	**0.4475**	
Root MSE	=	**.99725**	

(Std. Err. adjusted for **1,000** clusters in id)

y	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
x	**3.082817**	**.1056666**	**29.17**	**0.000**	**2.875463**	**3.290171**
_cons	**-.0447124**	**.0619673**	**-0.72**	**0.471**	**-.1663134**	**.0768886**

```
31 . eststo m3

32 .
33 . // Show all results
34 . esttab m*, se(3)
```

	(1) y	(2) y	(3) y
x	**3.083***	**3.083***	**3.083***
	(0.108)	**(0.054)**	**(0.106)**
_cons	**-0.0447**	**-0.0447**	**-0.0447**
	(0.062)	**(0.031)**	**(0.062)**
N	**1000**	**4000**	**4000**

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

```
35 .
   end of do-file

36 .
```