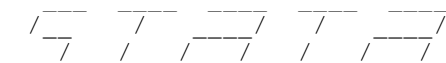
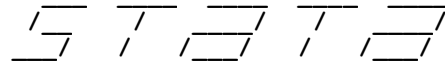


(R)

 Statistics/Data Analysis

User: Andong_Yan_HW1

(R)

 Statistics/Data Analysis

MP - Parallel Edition

14.0

Copyright 1985-2015 StataCorp LP
 StataCorp
 4905 Lakeway Drive
 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](#).

```
1 . doedit "C:\Users\yadto\Dropbox\USC\ECON513\HW2\hw2.do"
2 . do "C:\Users\yadto\Dropbox\USC\ECON513\HW2\hw2.do"
3 . log using c:\Users\yadto\Dropbox\USC\ECON513\HW2\hw2.log,replace
```

```
name: <unnamed>
log: c:\Users\yadto\Dropbox\USC\ECON513\HW2\hw2.log
log type: text
opened on: 22 Sep 2016, 15:59:01
```

```
4 .
5 . use c:\Users\yadto\Dropbox\USC\ECON513\HW2\hw2
6 .
7 . * Problem 1, question 1 *
8 . reg llearningswk educyears
```

Source	SS	df	MS	Number of obs	=	935
Model	20.0734039	1	20.0734039	F(1, 933)	=	115.09
Residual	162.726598	933	.174412216	Prob > F	=	0.0000
				R-squared	=	0.1098
				Adj R-squared	=	0.1089
Total	182.800002	934	.195717347	Root MSE	=	.41763

llearningswk	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educyears	.0667383	.0062209	10.73	0.000	.0545298	.0789469
_cons	5.045459	.0848918	59.43	0.000	4.878858	5.21206

```

9 . * the coefficient on education is the partial effect of education on the log weekly wage *
10 . * which is one year increase in education could bring how much percentage change in weekly wage
11 .
12 . * Problem 1, question 2 *
13 . * the likely sign of the omitted variable should be positive, so the bias should be upward *
14 . reg lnearningswk educyears iq

```

Source	SS	df	MS	Number of obs	=	935
Model	26.3200815	2	13.1600407	F(2, 932)	=	78.38
Residual	156.47992	932	.16789691	Prob > F	=	0.0000
				R-squared	=	0.1440
				Adj R-squared	=	0.1421
Total	182.800002	934	.195717347	Root MSE	=	.40975

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnearningswk						
educyears	.0443295	.007124	6.22	0.000	.0303486	.0583104
iq	.0063412	.0010396	6.10	0.000	.004301	.0083815
_cons	4.705016	.1002625	46.93	0.000	4.50825	4.901783

```

15 . * the coefficient on education decreases after adding IQ into regression *
16 . regress iq educyears

```

Source	SS	df	MS	Number of obs	=	935
Model	56280.9277	1	56280.9277	F(1, 933)	=	338.02
Residual	155346.531	933	166.502177	Prob > F	=	0.0000
				R-squared	=	0.2659
				Adj R-squared	=	0.2652
Total	211627.459	934	226.581862	Root MSE	=	12.904

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
iq						
educyears	3.533829	.1922095	18.39	0.000	3.156616	3.911042
_cons	53.68715	2.622933	20.47	0.000	48.53962	58.83469

```

17 . * we can see the regressors on education are both positive for the two regression equations *
18 . * according to the formula, we can see betahat = beta1 + beta2*cov(x1,x2)/var(x1) *
19 . * since IQ and education has a positive correlation, cov(x1,x2)>0, so betahat has a upward bias
20 .
21 . * Problem 2, question 5 *
22 . * Kx is the coefficient of x in regression of z on x, LamdaX is the coefficient of x in regression
> of z on x and w *
23 . reg iq educyears

```

Source	SS	df	MS	Number of obs	=	935
Model	56280.9277	1	56280.9277	F(1, 933)	=	338.02
Residual	155346.531	933	166.502177	Prob > F	=	0.0000
				R-squared	=	0.2659
				Adj R-squared	=	0.2652
Total	211627.459	934	226.581862	Root MSE	=	12.904

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
iq						
educyears	3.533829	.1922095	18.39	0.000	3.156616	3.911042
_cons	53.68715	2.622933	20.47	0.000	48.53962	58.83469

24 . reg iq educyears kww

Source	SS	df	MS	Number of obs	=	935
Model	67623.0292	2	33811.5146	F(2, 932)	=	218.83
Residual	144004.43	932	154.511191	Prob > F	=	0.0000
				R-squared	=	0.3195
				Adj R-squared	=	0.3181
Total	211627.459	934	226.581862	Root MSE	=	12.43

iq	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educyears	2.865716	.2009098	14.26	0.000	2.471428	3.260004
kww	.4949999	.0577748	8.57	0.000	.3816161	.6083836
_cons	44.99213	2.722911	16.52	0.000	39.64839	50.33588

25 . * so the proxy bias is smaller than the omitted variable bias, as the coefficient of x in second
> regression is smaller than the coefficient of x in first regression *

26 .

27 . log close
 name: <unnamed>
 log: c:\Users\yadto\Dropbox\USC\ECON513\HW2\hw2.log
 log type: text
 closed on: 22 Sep 2016, 15:59:01

28 .
 end of do-file

29 .