

MULTICOLLINEARITY WORKED PROBLEM

sum

Variable	Obs	Mean	Std. Dev.	Min	Max
age	3996	37.87813	13.29345	16	80
YRS_SCH	3996	12.95095	2.48391	0	18
earnwke	3996	421.3554	238.4387	0	999
schage	3996	18.95095	2.48391	6	24
exper	3996	18.92718	13.6901	-3	68

NOTE THAT EXPER = EXPERIENCE AND WAS CONSTRUCTED AS EXPER=AGE-YRS_SCH-6

```
. gen l_earn=ln( earnwke)
(1 missing value generated)
```

```
. reg l_earn age exper
```

Source	SS	df	MS	Number of obs = 3995		
Model	435.557047	2	217.778523	F(2, 3992)	=	495.01
Residual	1756.26068	3992	.439945061	Prob > F	=	0.0000
				R-squared	=	0.1987
				Adj R-squared	=	0.1983
				Root MSE	=	.66328
Total	2191.81773	3994	.548777599			

l_earn	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.1191409	.0043613	27.318	0.000	.1105903	.1276915
exper	-.1018068	.0042349	-24.040	0.000	-.1101096	-.0935041
_cons	3.249463	.0882069	36.839	0.000	3.076528	3.422397

```
. reg l_earn age gradeat
```

Source	SS	df	MS	Number of obs = 3995		
Model	435.557047	2	217.778523	F(2, 3992)	=	495.01
Residual	1756.26068	3992	.439945061	Prob > F	=	0.0000
				R-squared	=	0.1987
				Adj R-squared	=	0.1983
				Root MSE	=	.66328
Total	2191.81773	3994	.548777599			

l_earn	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0173341	.0007913	21.904	0.000	.0157826	.0188856
YRS_SCH	.1018068	.0042349	24.040	0.000	.0935041	.1101096
_cons	3.860304	.0651376	59.264	0.000	3.732598	3.98801

```
. reg l_earn age exper if gradeat=12
invalid syntax
r(198);
```

```
. reg l_earn age exper if gradeat==12
```

Source	SS	df	MS	Number of obs = 1890		
Model	54.2083669	1	54.2083669	F(1, 1888)	=	149.11
Residual	686.371371	1888	.363544159	Prob > F	=	0.0000
				R-squared	=	0.0732
				Adj R-squared	=	0.0727
				Root MSE	=	.60295
Total	740.579738	1889	.392048564			

l_earn	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0132565	.0010856	12.211	0.000	.0111274	.0153856
exper	(dropped)					
_cons	5.294551	.042518	124.525	0.000	5.211163	5.377938

** PROBLEM OF MULTICOLLINEARITY NOTED**

```
. reg exper age
```

Source	SS	df	MS	Number of obs =	3996
Model	724205.56	1	724205.56	F(1, 3994) =	.
Residual	24532.2485	3994	6.14227554	Prob > F =	0.0000
				R-squared =	0.9672
				Adj R-squared =	0.9672
Total	748737.809	3995	187.418726	Root MSE =	2.4784

exper	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	1.012826	.0029496	343.373	0.000	1.007043	1.018609
_cons	-19.43678	.1184059	-164.154	0.000	-19.66892	-19.20463

GENERATE DUMMY VARIABLE FOR AN AGE PROFILE - NO LONGER COLINEAR WITH EXPERIENCE

```
. gen b50_60=0

. recode b50_60 0=1 if age<46&age>34
(1051 changes made)

. gen b61_70=0

. recode b61_70 0=1 if age<35&age>24
(1142 changes made)

. reg l_earn b50_60 b61_70 exper if gradeat==12
```

Source	SS	df	MS	Number of obs =	1890
Model	125.787306	3	41.9291019	F(3, 1886) =	128.63
Residual	614.792432	1886	.3259769	Prob > F =	0.0000
				R-squared =	0.1698
				Adj R-squared =	0.1685
Total	740.579738	1889	.392048564	Root MSE =	.57094

l_earn	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
b50_60	.4173457	.0324998	12.841	0.000	.3536064	.4810851
b61_70	.3903027	.0335125	11.646	0.000	.3245773	.4560281
exper	.0169014	.0011177	15.122	0.000	.0147093	.0190934
_cons	5.237259	.0321462	162.920	0.000	5.174213	5.300305