

Stata Textbook Examples**Introductory Econometrics: A Modern Approach by Jeffrey M. Wooldridge (1st & 2nd eds.)****Chapter 7 - Multiple Regression Analysis with Qualitative Information: Binary (or Dummy) Variables****Example 7.1: Hourly Wage Equation**

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE1
```

```
reg wage female educ exper tenure
```

Source	SS	df	MS	Number of obs =	526
Model	2603.10658	4	650.776644	F(4, 521) =	74.40
Residual	4557.30771	521	8.7472317	Prob > F =	0.0000
				R-squared =	0.3635
				Adj R-squared =	0.3587
Total	7160.41429	525	13.6388844	Root MSE =	2.9576

wage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
female	-1.810852	.2648252	-6.84	0.000	-2.331109	-1.290596
educ	.5715048	.0493373	11.58	0.000	.4745803	.6684293
exper	.0253959	.0115694	2.20	0.029	.0026674	.0481243
tenure	.1410051	.0211617	6.66	0.000	.0994323	.1825778
_cons	-1.567939	.7245511	-2.16	0.031	-2.991339	-.144538

```
reg wage female
```

Source	SS	df	MS	Number of obs =	526
Model	828.220467	1	828.220467	F(1, 524) =	68.54
Residual	6332.19382	524	12.0843394	Prob > F =	0.0000
				R-squared =	0.1157
				Adj R-squared =	0.1140
Total	7160.41429	525	13.6388844	Root MSE =	3.4763

wage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
female	-2.51183	.3034092	-8.28	0.000	-3.107878	-1.915782
_cons	7.099489	.2100082	33.81	0.000	6.686928	7.51205

Average wage for women

```
lincom female+_cons
```

```
( 1) female + _cons = 0.0
```

wage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	4.587659	.2189834	20.95	0.000	4.157466	5.017852

Example 7.2: Effects of Computer Ownership on College GPA

use <http://fmwww.bc.edu/ec-p/data/wooldridge/GPA1>

reg colGPA PC hsGPA ACT

Source	SS	df	MS	Number of obs = 141		
Model	4.25741863	3	1.41913954	F(3, 137) = 12.83		
Residual	15.1486808	137	.110574313	Prob > F = 0.0000		
Total	19.4060994	140	.138614996	R-squared = 0.2194		
				Adj R-squared = 0.2023		
				Root MSE = .33253		

colGPA	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
PC	.1573092	.0572875	2.75	0.007	.0440271	.2705913
hsGPA	.4472417	.0936475	4.78	0.000	.2620603	.632423
ACT	.008659	.0105342	0.82	0.413	-.0121717	.0294897
_cons	1.26352	.3331255	3.79	0.000	.6047871	1.922253

reg colGPA PC

Source	SS	df	MS	Number of obs = 141		
Model	.970092892	1	.970092892	F(1, 139) = 7.31		
Residual	18.4360066	139	.132633141	Prob > F = 0.0077		
Total	19.4060994	140	.138614996	R-squared = 0.0500		
				Adj R-squared = 0.0432		
				Root MSE = .36419		

colGPA	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
PC	.1695168	.0626805	2.70	0.008	.0455864	.2934472
_cons	2.989412	.0395018	75.68	0.000	2.91131	3.067514

Example 7.3: Effects of Training Grants on Hours of Training in 1988

use <http://fmwww.bc.edu/ec-p/data/wooldridge/JTRAIN>

reg hrsemp grant lsales lemploy if year==1988

Source	SS	df	MS	Number of obs = 105			
Model	18622.7243	3	6207.57476	F(3, 101)	=	10.44	
Residual	60031.0957	101	594.367284	Prob > F	=	0.0000	
				R-squared	=	0.2368	
				Adj R-squared	=	0.2141	
Total	78653.82	104	756.286731	Root MSE	=	24.38	

hrsemp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
grant	26.2545	5.591766	4.70	0.000	15.16194	37.34706
lsales	-.9845776	3.539904	-0.28	0.781	-8.006795	6.03764
lemploy	-6.069873	3.882894	-1.56	0.121	-13.77249	1.632744
_cons	46.66504	43.41211	1.07	0.285	-39.4529	132.783

Example 7.4: Housing Price Regression

use <http://fmwww.bc.edu/ec-p/data/wooldridge/HPRICE1>

reg lprice llotsize lsqrft bdrms colonial

Source	SS	df	MS	Number of obs = 88			
Model	5.20400088	4	1.30100022	F(4, 83)	=	38.38	
Residual	2.81362108	83	.033899049	Prob > F	=	0.0000	
				R-squared	=	0.6491	
				Adj R-squared	=	0.6322	
Total	8.01762195	87	.092156574	Root MSE	=	.18412	

lprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
llotsize	.1678202	.0381806	4.40	0.000	.0918805	.2437599
lsqrft	.7071932	.0928019	7.62	0.000	.5226139	.8917725
bdrms	.0268308	.0287235	0.93	0.353	-.0302992	.0839608
colonial	.0537949	.0447732	1.20	0.233	-.0352572	.142847
_cons	5.558154	.6510406	8.54	0.000	4.263261	6.853048

Example 7.5: Log Hourly Wage Equation

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE1
```

```
reg lwage female educ exper expersq tenure tenursq
```

Source	SS	df	MS	Number of obs = 526			
Model	65.3791002	6	10.8965167	F(6, 519) = 68.18			
Residual	82.9506616	519	.159827864	Prob > F = 0.0000			
				R-squared = 0.4408			
				Adj R-squared = 0.4343			
Total	148.329762	525	.28253288	Root MSE = .39978			

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
female	-.296511	.0358055	-8.28	0.000	-.3668524	-.2261695
educ	.0801967	.0067573	11.87	0.000	.0669217	.0934716
exper	.0294324	.0049752	5.92	0.000	.0196584	.0392063
expersq	-.0005827	.0001073	-5.43	0.000	-.0007935	-.0003719
tenure	.0317139	.0068452	4.63	0.000	.0182663	.0451616
tenursq	-.0005852	.0002347	-2.49	0.013	-.0010463	-.0001241
_cons	.4166909	.0989279	4.21	0.000	.2223425	.6110393

Difference between woman's and man's wage

```
di exp(_b[female]*1)-1
-.25659254
```

Example 7.6: Log Hourly Wage Equation

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE1
```

```
gen male = (!female)
```

```
gen single = (~married)
```

```
gen marrmale = (married & ~female)
```

```
gen marrfem = (married & female)
```

```
gen singfem = (female & ~married)
```

```
gen singmale = (~female & ~married)
```

```
reg lwage marrmale marrfem singfem educ exper expersq tenure tenursq
```

Source	SS	df	MS	Number of obs = 526			
--------	----	----	----	---------------------	--	--	--

-----+-----				F(8, 517) = 55.25
Model	68.3617614	8	8.54522017	Prob > F = 0.0000
Residual	79.9680004	517	.154676983	R-squared = 0.4609
-----+-----				Adj R-squared = 0.4525
Total	148.329762	525	.28253288	Root MSE = .39329

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
marrmale	.2126756	.0553572	3.84	0.000	.103923	.3214283
marrfem	-.1982676	.0578355	-3.43	0.001	-.3118891	-.0846462
singfem	-.1103502	.0557421	-1.98	0.048	-.219859	-.0008414
educ	.0789103	.0066945	11.79	0.000	.0657585	.0920621
exper	.0268006	.0052428	5.11	0.000	.0165007	.0371005
expersq	-.0005352	.0001104	-4.85	0.000	-.0007522	-.0003183
tenure	.0290875	.006762	4.30	0.000	.0158031	.0423719
tenursq	-.0005331	.0002312	-2.31	0.022	-.0009874	-.0000789
_cons	.321378	.100009	3.21	0.001	.1249041	.517852

Difference in lwage between married and single women

lincom singfem-marrfem

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	.0879174	.0523481	1.68	0.094	-.0149238	.1907587

reg lwage marrmale singmale singfem educ exper expersq tenure tenursq

Source	SS	df	MS	Number of obs = 526	
-----+-----				F(8, 517) = 55.25	
Model	68.3617614	8	8.54522017	Prob > F = 0.0000	
Residual	79.9680004	517	.154676983	R-squared = 0.4609	
-----+-----				Adj R-squared = 0.4525	
Total	148.329762	525	.28253288	Root MSE = .39329	

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
marrmale	.4109433	.0457709	8.98	0.000	.3210234	.5008631
singmale	.1982676	.0578355	3.43	0.001	.0846462	.3118891
singfem	.0879174	.0523481	1.68	0.094	-.0149238	.1907587
educ	.0789103	.0066945	11.79	0.000	.0657585	.0920621
exper	.0268006	.0052428	5.11	0.000	.0165007	.0371005
expersq	-.0005352	.0001104	-4.85	0.000	-.0007522	-.0003183
tenure	.0290875	.006762	4.30	0.000	.0158031	.0423719
tenursq	-.0005331	.0002312	-2.31	0.022	-.0009874	-.0000789

_cons		.1231104	.1057937	1.16	0.245	-.084728	.3309488
-------	--	----------	----------	------	-------	----------	----------

Example 7.7: Effects of Physical Attractiveness on Wage

Dataset is not available

Example 7.8: Effects of Law School Rankings on Starting Salaries

use <http://fmwww.bc.edu/ec-p/data/wooldridge/LAWSCH85>

gen r61_100 = (rank>60 & rank<101)

reg lsalary top10 r11_25 r26_40 r41_60 r61_100 LSAT GPA llibvol lcost

Source	SS	df	MS	Number of obs =	136
Model	9.45225307	9	1.05025034	F(9, 126) =	143.20
Residual	.924109594	126	.007334203	Prob > F =	0.0000
				R-squared =	0.9109
				Adj R-squared =	0.9046
Total	10.3763627	135	.076861946	Root MSE =	.08564

lsalary	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
top10	.6995646	.0534919	13.08	0.000	.5937057	.8054236
r11_25	.5935444	.03944	15.05	0.000	.5154938	.6715951
r26_40	.3750779	.0340812	11.01	0.000	.3076322	.4425236
r41_60	.26282	.027962	9.40	0.000	.2074839	.3181561
r61_100	.1315946	.0210418	6.25	0.000	.0899535	.1732358
LSAT	.0056908	.003063	1.86	0.066	-.0003708	.0117524
GPA	.0137274	.0741919	0.19	0.854	-.1330962	.1605509
llibvol	.0363614	.0260165	1.40	0.165	-.0151245	.0878472
lcost	.0008418	.025136	0.03	0.973	-.0489017	.0505852
_cons	9.165292	.4114241	22.28	0.000	8.351096	9.979488

Difference in starting wage between top 10 below 100 school

*di exp(_[top10]*1)-1*

1.0137

reg lsalary rank LSAT GPA llibvol lcost

Source		SS	df	MS	Number of obs =	136
--------	--	----	----	----	-----------------	-----

-----+-----				F(5, 130) = 138.23		
Model		8.73363382	5	1.74672676	Prob > F = 0.0000	
Residual		1.64272884	130	.012636376	R-squared = 0.8417	
-----+-----				Adj R-squared = 0.8356		
Total		10.3763627	135	.076861946	Root MSE = .11241	

lsalary		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
-----+-----						
rank		-.0033246	.0003485	-9.54	0.000	-.004014 -.0026352
LSAT		.0046964	.0040105	1.17	0.244	-.0032379 .0126307
GPA		.2475245	.090037	2.75	0.007	.069397 .4256519
llibvol		.0949925	.0332543	2.86	0.005	.0292028 .1607823
lcost		.0375543	.0321061	1.17	0.244	-.0259637 .1010723
_cons		8.343234	.5325191	15.67	0.000	7.289709 9.396759

Example 7.9: Effects of Computer Usage on Wages

Dataset is not available

Example 7.10: Log Hourly Wage Equation

use <http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE2>

*gen femed = female*educ*

reg lwage female educ femed exper expersq tenure tenursq

Source		SS	df	MS	Number of obs = 526		
-----+-----					F(7, 518) = 58.37		
Model		65.4081526	7	9.3440218	Prob > F = 0.0000		
Residual		82.9216091	518	.160080326	R-squared = 0.4410		
-----+-----					Adj R-squared = 0.4334		
Total		148.329762	525	.28253288	Root MSE = .4001		

lwage		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----+-----							
female		-.2267887	.1675394	-1.35	0.176	-.555929	.1023516
educ		.0823692	.0084699	9.72	0.000	.0657296	.0990088
femed		-.0055645	.0130618	-0.43	0.670	-.0312252	.0200962
exper		.0293366	.0049842	5.89	0.000	.019545	.0391283
expersq		-.0005804	.0001075	-5.40	0.000	-.0007916	-.0003691
tenure		.0318967	.006864	4.65	0.000	.018412	.0453814

tenursq		-.00059	.0002352	-2.51	0.012	-.001052	-.000128
_cons		.388806	.1186871	3.28	0.001	.1556388	.6219733

```
reg lwage female educ exper expersq tenure tenursq
```

Source	SS	df	MS	Number of obs = 526			
Model	65.3791002	6	10.8965167	F(6, 519) = 68.18			
Residual	82.9506616	519	.159827864	Prob > F = 0.0000			
Total	148.329762	525	.28253288	R-squared = 0.4408			
				Adj R-squared = 0.4343			
				Root MSE = .39978			

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
female	-.296511	.0358055	-8.28	0.000	-.3668524	-.2261695
educ	.0801967	.0067573	11.87	0.000	.0669217	.0934716
exper	.0294324	.0049752	5.92	0.000	.0196584	.0392063
expersq	-.0005827	.0001073	-5.43	0.000	-.0007935	-.0003719
tenure	.0317139	.0068452	4.63	0.000	.0182663	.0451616
tenursq	-.0005852	.0002347	-2.49	0.013	-.0010463	-.0001241
_cons	.4166909	.0989279	4.21	0.000	.2223425	.6110393

Example 7.11: Effects of Race on Baseball Player Salaries

use <http://fmwww.bc.edu/ec-p/data/wooldridge/MLB1>

```
reg lsalary years gamesyr bavg hrunsyr rbisyr runsyr fldperc allstar black hispan
blkpbb hispbb
```

Source	SS	df	MS	Number of obs = 330			
Model	283.782211	12	23.6485176	F(12, 317) = 46.48			
Residual	161.279291	317	.50876748	Prob > F = 0.0000			
Total	445.061503	329	1.35277053	R-squared = 0.6376			
				Adj R-squared = 0.6239			
				Root MSE = .71328			

lsalary	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
years	.0673458	.0128915	5.22	0.000	.0419821	.0927094
gamesyr	.0088778	.0033837	2.62	0.009	.0022205	.0155352
bavg	.0009451	.0015133	0.62	0.533	-.0020322	.0039225
hrunsyr	.0146206	.0164522	0.89	0.375	-.0177488	.04699
rbisyr	.0044938	.007575	0.59	0.553	-.0104098	.0193974
runsyr	.0072029	.0045671	1.58	0.116	-.0017827	.0161884

fldperc		.0010865	.0021195	0.51	0.609	-.0030836	.0052566
allstar		.0075307	.0028735	2.62	0.009	.0018771	.0131843
black		-.1980075	.1254968	-1.58	0.116	-.4449192	.0489043
hispan		-.1900079	.1530902	-1.24	0.215	-.491209	.1111933
blckpb		.0124513	.0049628	2.51	0.013	.0026871	.0222154
hispph		.0200862	.0097933	2.05	0.041	.0008181	.0393543
_cons		10.34369	2.182538	4.74	0.000	6.0496	14.63778

Difference in lwage between black and white in cities with 10% of blacks

```
lincom _b[black]+_b[blckpb]*10
```

```
( 1) black + 10.0 blckpb = 0.0
```

lsalary		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
(1)		-.0734949	.0997916	-0.74	0.462	-.2698324 .1228426

Difference in lwage between black and white in cities with 20% of blacks

```
lincom _b[black]+_b[blckpb]*20
```

```
( 1) black + 20.0 blckpb = 0.0
```

lsalary		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
(1)		.0510177	.0953577	0.54	0.593	-.1365962 .2386316

City percentage of hispanic people when wages of hispanic and whites are equal

```
di _b[hispan]*-1/_b[hispph]
9.4596276
```

Example 7.12: A Linear Probability Model of Arrests

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/CRIME1
```

```
gen arr86=(~narr86)
```

```
reg arr86 pcnv avgsen tottime ptime86 qemp86
```

Source		SS	df	MS	Number of obs =	2725
--------	--	----	----	----	-----------------	------

Model	25.8452455	5	5.16904909	F(5, 2719) =	27.03
Residual	519.971268	2719	.191236215	Prob > F =	0.0000
				R-squared =	0.0474
				Adj R-squared =	0.0456
Total	545.816514	2724	.20037317	Root MSE =	.43731

arr86	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
pcnv	.1624448	.0212368	7.65	0.000	.120803	.2040866
avgsen	-.0061127	.006452	-0.95	0.344	-.018764	.0065385
totttime	.0022616	.0049781	0.45	0.650	-.0074997	.0120229
ptime86	.0219664	.0046349	4.74	0.000	.0128781	.0310547
qemp86	.0428294	.0054046	7.92	0.000	.0322319	.0534268
_cons	.5593846	.0172329	32.46	0.000	.5255937	.5931754

Change in probability of arrest if pcnv increases by .5

lincom _b[pcnv]*.5

(1) .5 pcnv = 0.0

arr86	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	.0812224	.0106184	7.65	0.000	.0604015	.1020433

Change in probability of arrest if ptime86 increases by 6

lincom _b[ptime86]*6

(1) 6.0 ptime86 = 0.0

arr86	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	.1317984	.0278095	4.74	0.000	.0772686	.1863282

Change in probability of arrest if ptime86 decreases by 12

lincom _b[_cons]- _b[ptime86]*12

(1) - 12.0 ptime86 + _cons = 0.0

arr86	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-------	-------	-----------	---	------	----------------------	--

```

-----+-----
(1) | .2957878 .061983 4.77 0.000 .1742492 .4173264
-----+-----

```

Change in probability of arrest if qemp86 increases by 4

```
lincom _b[qemp86]*4
```

```
( 1) 4.0 qemp86 = 0.0
```

```

-----+-----
arr86 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----
(1)   | .1713175   .0216182     7.92   0.000   .1289277   .2137073
-----+-----

```

```
reg arr86 pcnv avgsen tottime ptime86 qemp86 black hispan
```

```

-----+-----
Source |      SS      df      MS                Number of obs =      2725
-----+-----
Model  | 37.2205275      7   5.31721822           F( 7, 2717) =      28.41
Residual | 508.595986    2717   .187190278         Prob > F      =      0.0000
-----+-----
Total   | 545.816514    2724   .20037317          R-squared      =      0.0682
                                           Adj R-squared  =      0.0658
                                           Root MSE      =      .43265

```

```

-----+-----
arr86 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----
pcnv  | .152062   .0210655     7.22   0.000   .1107561   .193368
avgsen | -.0046191 .0063888    -0.72   0.470   -.0171465   .0079083
tottime | .0025619 .0049259     0.52   0.603   -.0070969   .0122207
ptime86 | .0236954 .0045948     5.16   0.000   .0146858   .032705
qemp86 | .0384737 .0054016     7.12   0.000   .0278821   .0490653
black  | -.1697631 .0236738    -7.17   0.000   -.2161836  -.1233426
hispan | -.0961866 .0207105    -4.64   0.000   -.1367965  -.0555766
_cons  | .6195717 .0187272    33.08   0.000   .5828507   .6562927
-----+-----

```

This page prepared by Oleksandr Talavera (revised 8 Nov 2002)

Send your questions/comments/suggestions to Kit Baum at baum@bc.edu
 These pages are maintained by the Faculty Micro Resource Center's **GSA Program**,
 a unit of Boston College **Academic Technology Services**