Stata Textbook Examples
Introductory Econometrics: A Modern Approach by Jeffrey M. Wooldridge (1st & 2nd eds.)
Chapter 6 - Multiple Regression Analysis: Further Issues

Example 6.1: Effect of Pollution on Housing Prices

use *http://fmwww.bc.edu/ec-p/data/wooldridge/HPRICE2*

reg price nox crime rooms dist stratio, beta

```
  Source |       SS       df       MS                  Number of obs =     506
---------+------------------------------               F(  5,   500) =  174.47
   Model |  2.7223e+10      5   5.4445e+09             Prob > F      =  0.0000
Residual |  1.5603e+10    500   31205611.6             R-squared     =  0.6357
---------+------------------------------               Adj R-squared =  0.6320
   Total |  4.2826e+10    505   84803032.0             Root MSE      =  5586.2
```

```
-------------------------------------------------------------------------------------
   price |      Coef.   Std. Err.        t    P>|t|                             Beta
---------+---------------------------------------------------------------------------
     nox | -2706.433   354.0869     -7.643   0.000                        -.340446
   crime |  -153.601   32.92883     -4.665   0.000                       -.1432828
   rooms |  6735.498   393.6037     17.112   0.000                        .5138878
    dist | -1026.806   188.1079     -5.459   0.000                       -.2348385
 stratio | -1149.204   127.4287     -9.018   0.000                       -.2702799
   _cons |  20871.13   5054.599      4.129   0.000                               .
-------------------------------------------------------------------------------------
```

Example 6.2: Effect of Pollution on Housing Prices

use *http://fmwww.bc.edu/ec-p/data/wooldridge/PRICE2*

gen *rooms2=rooms*rooms*

gen *ldist=log(dist)*

reg lprice lnox ldist rooms rooms2 stratio

```
  Source |       SS       df       MS                  Number of obs =     506
---------+------------------------------               F(  5,   500) =  151.77
   Model |  50.98725       5    10.19745              Prob > F      =  0.0000
Residual | 33.595021     500   .067190042             R-squared     =  0.6028
---------+------------------------------               Adj R-squared =  0.5988
   Total | 84.5822709    505   .167489645             Root MSE      =  .25921
```

```
   lprice |      Coef.   Std. Err.        t    P>|t|     [95% Conf. Interval]
---------+----------------------------------------------------------------------
     lnox | -.9016832     .114687     -7.862   0.000    -1.127011    -.6763553
    ldist | -.0867821    .0432808     -2.005   0.045     -.1718166   -.0017475
    rooms | -.5451122    .1654542     -3.295   0.001     -.8701834    -.220041
   rooms2 |  .0622611     .012805      4.862   0.000      .0371029    .0874194
  stratio | -.0475903    .0058542     -8.129   0.000     -.0590921   -.0360884
    _cons |  13.38548    .5664734     23.629   0.000      12.27252    14.49844
```

Turnaround value of rooms

*display -1*_b[rooms]/(2*_b[rooms2])*
4.3776278

Change in price if rooms increases from 5 to 6

display 100(_b[rooms]+2*_b[rooms2]*5)*
7.7499207

Change in price if rooms increases from 6 to 7

display 100(_b[rooms]+2*_b[rooms2]*6)*
20.202149

Example 6.3: Effect of Attendance on Final Exam Performance

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

summ priGPA
```
Variable |     Obs      Mean   Std. Dev.       Min        Max
---------+-----------------------------------------------------
  priGPA |     680   2.586775   .5447141      .857       3.93
```

*gen priGPA2=priGPA*priGPA*

*gen ACT2=ACT*ACT*

*gen priatn=priGPA*atndrte*

reg stndfnl atndrte priGPA ACT priGPA2 ACT2 priatn
```
  Source |       SS       df       MS                   Number of obs =     680
```

```
--------+------------------------------        F(  6,   673) =    33.25
  Model |  152.001001     6  25.3335002        Prob > F      =   0.0000
Residual |   512.76244   673  .761905557       R-squared     =   0.2287
--------+------------------------------        Adj R-squared =   0.2218
  Total |  664.763441   679   .97903305        Root MSE      =  .87287
```

```
------------------------------------------------------------------------
 stndfnl |    Coef.    Std. Err.      t     P>|t|    [95% Conf. Interval]
--------+---------------------------------------------------------------
 atndrte | -.0067129    .0102321   -0.656   0.512   -.0268035    .0133777
  priGPA | -1.62854     .4810025   -3.386   0.001   -2.572986   -.6840938
     ACT | -.1280394    .098492    -1.300   0.194   -.3214279    .0653492
 priGPA2 |  .2959046    .1010495    2.928   0.004    .0974945    .4943147
    ACT2 |  .0045334    .0021764    2.083   0.038     .00026     .0088068
  priatn |  .0055859    .0043174    1.294   0.196   -.0028913    .0140631
   _cons |  2.050293   1.360319     1.507   0.132   -.6206864   4.721272
------------------------------------------------------------------------
```

Partial effect of atndrte on stndfnl

display _b[atndrte]+_b[priatn]*2.59
.00775457

Example 6.4: CEO Compensation and Firm Performance

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

reg salary sales roe

```
 Source |     SS       df      MS             Number of obs =     209
--------+------------------------------        F(  2,   206) =     3.09
  Model |  11427511.8     2  5713755.89        Prob > F      =  0.0474
Residual |  380305470   206  1846143.06        R-squared     =  0.0292
--------+------------------------------        Adj R-squared =  0.0197
  Total |  391732982   208  1883331.64         Root MSE      =  1358.7
```

```
------------------------------------------------------------------------
 salary  |    Coef.    Std. Err.      t     P>|t|    [95% Conf. Interval]
--------+---------------------------------------------------------------
  sales  |  .0163416    .0088736    1.842   0.067   -.0011532    .0338363
    roe  | 19.63097    11.07655     1.772   0.078   -2.20697    41.46891
  _cons  | 830.6313    223.9049     3.710   0.000   389.1924    1272.07
------------------------------------------------------------------------
```

reg lsalary lsales roe

```
  Source |       SS       df       MS                    Number of obs =      209
---------+------------------------------                 F(  2,    206) =    40.45
   Model | 18.8149023      2  9.40745113                 Prob > F      =   0.0000
Residual | 47.9072676    206  .232559552                 R-squared     =   0.2820
---------+------------------------------                 Adj R-squared =   0.2750
   Total | 66.7221699    208  .320779663                 Root MSE      =   .48224


-----------------------------------------------------------------------------------
 lsalary |      Coef.   Std. Err.       t     P>|t|      [95% Conf. Interval]
---------+-------------------------------------------------------------------------
  lsales |   .2750875    .033254      8.272   0.000      .2095258    .3406492
     roe |   .0178723   .0039551      4.519   0.000      .0100746    .0256699
   _cons |   4.362167   .2938776     14.843   0.000      3.782774    4.941561
-----------------------------------------------------------------------------------
```

Example 6.5: Confidence Interval for Predicted College GPA (Approach in Book)

use http://fmwww.bc.edu/ec-p/data/wooldridge/GPA2

*gen hsize2=hsize*hsize*

reg colgpa sat hsperc hsize hsize2

```
      Source |       SS       df       MS                   Number of obs =     4137
------------+------------------------------                 F(  4,   4132) =   398.02
       Model | 499.030504      4  124.757626                Prob > F      =   0.0000
    Residual | 1295.16517   4132  .313447524                R-squared     =   0.2781
------------+------------------------------                 Adj R-squared =   0.2774
       Total | 1794.19567   4136  .433799728                Root MSE      =   .55986


-----------------------------------------------------------------------------------
     colgpa |      Coef.   Std. Err.       t     P>|t|      [95% Conf. Interval]
------------+----------------------------------------------------------------------
        sat |   .0014925   .0000652     22.89   0.000      .0013646    .0016204
     hsperc |  -.0138558    .000561    -24.70   0.000     -.0149557   -.0127559
      hsize |  -.0608815   .0165012     -3.69   0.000     -.0932327   -.0285302
     hsize2 |   .0054603   .0022698      2.41   0.016      .0010102    .0099104
      _cons |   1.492652   .0753414     19.81   0.000      1.344942    1.640362
-----------------------------------------------------------------------------------
```

Predicted college GPA

*display _b[_cons]+_b[sat]*1200+_b[hsperc]*30+_b[hsize]*5+_b[hsize2]*25*
```
2.7000755
```

```
gen sat0=sat-1200
```

```
gen hsperc0=hsperc-30
```

```
gen hsize0=hsize-5
```

```
gen hsize20=hsize2-25
```

```
reg colgpa sat0 hsperc0 hsize0 hsize20
```

```
      Source |       SS       df       MS                  Number of obs =     4137
-------------+------------------------------              F(  4,  4132) =  398.02
       Model | 499.030503        4  124.757626            Prob > F      =  0.0000
    Residual | 1295.16517     4132  .313447524            R-squared     =  0.2781
-------------+------------------------------              Adj R-squared =  0.2774
       Total | 1794.19567     4136  .433799728            Root MSE      =  .55986


------------------------------------------------------------------------------
      colgpa |      Coef.   Std. Err.        t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        sat0 |   .0014925   .0000652     22.89    0.000     .0013646    .0016204
     hsperc0 |  -.0138558    .000561    -24.70    0.000    -.0149557   -.0127559
      hsize0 |  -.0608815   .0165012     -3.69    0.000    -.0932327   -.0285302
     hsize20 |   .0054603   .0022698      2.41    0.016     .0010102    .0099104
       _cons |   2.700075   .0198778    135.83    0.000     2.661104    2.739047
------------------------------------------------------------------------------
```

Example 6.5: Confidence Interval for Predicted College GPA (Another Approach)

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/GPA2
```

```
gen hsize2=hsize*hsize
```

```
reg colgpa sat hsperc hsize hsize2
```

```
      Source |       SS       df       MS                  Number of obs =     4137
-------------+------------------------------              F(  4,  4132) =  398.02
       Model | 499.030504        4  124.757626            Prob > F      =  0.0000
    Residual | 1295.16517     4132  .313447524            R-squared     =  0.2781
-------------+------------------------------              Adj R-squared =  0.2774
       Total | 1794.19567     4136  .433799728            Root MSE      =  .55986


------------------------------------------------------------------------------
      colgpa |      Coef.   Std. Err.        t    P>|t|     [95% Conf. Interval]
```

```
------------+----------------------------------------------------------------
        sat |   .0014925   .0000652    22.89   0.000     .0013646    .0016204
     hsperc |  -.0138558    .000561   -24.70   0.000    -.0149557   -.0127559
      hsize |  -.0608815   .0165012    -3.69   0.000    -.0932327   -.0285302
     hsize2 |   .0054603   .0022698     2.41   0.016     .0010102    .0099104
      _cons |   1.492652   .0753414    19.81   0.000     1.344942    1.640362
------------------------------------------------------------------------------
```

set obs 4138

replace sat=1200 in 4138/4138

replace hsperc=30 in 4138/4138

replace hsize=5 in 4138/4138

replace hsize2=25 in 4138/4138

regress

```
      Source |       SS       df       MS                  Number of obs =    4137
-------------+------------------------------              F(  4,  4132) =  398.02
       Model |  499.030504     4  124.757626              Prob > F      =  0.0000
    Residual |  1295.16517  4132  .313447524              R-squared     =  0.2781
-------------+------------------------------              Adj R-squared =  0.2774
       Total |  1794.19567  4136  .433799728              Root MSE      =  .55986


------------------------------------------------------------------------------
     colgpa |      Coef.   Std. Err.       t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        sat |   .0014925   .0000652    22.89   0.000     .0013646    .0016204
     hsperc |  -.0138558    .000561   -24.70   0.000    -.0149557   -.0127559
      hsize |  -.0608815   .0165012    -3.69   0.000    -.0932327   -.0285302
     hsize2 |   .0054603   .0022698     2.41   0.016     .0010102    .0099104
      _cons |   1.492652   .0753414    19.81   0.000     1.344942    1.640362
------------------------------------------------------------------------------
```

predict colgpahat in 4138/4138,stdp

predict colgpahatt in 4138/4138,xb

gen lb = colgpahatt-1.96* colgpahat in 4138/4138

gen ub = colgpahatt+1.96* colgpahat in 4138/4138

list colgpahat lb colgpahatt ub in 4138/4138

```
        colgpahat          lb  colgpahatt          ub
4138.    .0198778    2.661115    2.700075    2.739036
```

Example 6.6: Confidence Interval for Future College GPA

use http://fmwww.bc.edu/ec-p/data/wooldridge/GPA2

gen hsize2=hsize*hsize

reg colgpa sat hsperc hsize hsize2

```
      Source |       SS       df       MS              Number of obs =    4137
-------------+------------------------------           F(  4,  4132) =  398.02
       Model | 499.030504        4  124.757626          Prob > F      =  0.0000
    Residual | 1295.16517     4132  .313447524          R-squared     =  0.2781
-------------+------------------------------           Adj R-squared =  0.2774
       Total | 1794.19567     4136  .433799728          Root MSE      =  .55986


------------------------------------------------------------------------------
      colgpa |      Coef.   Std. Err.       t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sat |   .0014925   .0000652     22.89   0.000     .0013646    .0016204
      hsperc |  -.0138558    .000561    -24.70   0.000    -.0149557   -.0127559
       hsize |  -.0608815   .0165012     -3.69   0.000    -.0932327   -.0285302
      hsize2 |   .0054603   .0022698      2.41   0.016     .0010102    .0099104
       _cons |   1.492652   .0753414     19.81   0.000     1.344942    1.640362
------------------------------------------------------------------------------
```

set obs 4138

replace sat=1200 in 4138/4138

replace hsperc=30 in 4138/4138

replace hsize=5 in 4138/4138

replace hsize2=25 in 4138/4138

regress

```
      Source |       SS       df       MS              Number of obs =    4137
-------------+------------------------------           F(  4,  4132) =  398.02
       Model | 499.030504        4  124.757626          Prob > F      =  0.0000
```

```
      Residual |  1295.16517    4132  .313447524           R-squared     =  0.2781
---------------+------------------------------           Adj R-squared =  0.2774
         Total |  1794.19567    4136  .433799728           Root MSE      =  .55986


-------------------------------------------------------------------------------
        colgpa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+---------------------------------------------------------------
           sat |   .0014925   .0000652     22.89   0.000     .0013646    .0016204
        hsperc |  -.0138558    .000561    -24.70   0.000    -.0149557   -.0127559
         hsize |  -.0608815   .0165012     -3.69   0.000    -.0932327   -.0285302
        hsize2 |   .0054603   .0022698      2.41   0.016     .0010102    .0099104
         _cons |   1.492652   .0753414     19.81   0.000     1.344942    1.640362
-------------------------------------------------------------------------------
```

predict cc in 4138/4138,stdf

predict colgpahatt in 4138/4138,xb

gen lb1 = colgpahatt-1.96 cc in 4138/4138*

gen ub1 = colgpahatt+1.96 cc in 4138/4138*

list cc lb1 colgpahatt ub1 in 4138/4138

```
               cc        lb1   colgpahatt         ub1
4138.   .5602166   1.602051    2.700075      3.7981
```

Example 6.7: Predicting CEO Salaries

use http://fmwww.bc.edu/ec-p/data/wooldridge/CEOSAL2

reg lsalary lsales lmktval ceoten

```
      Source |       SS       df       MS              Number of obs =     177
-------------+------------------------------           F(  3,   173) =   26.91
       Model |  20.5672427       3  6.85574758           Prob > F      =  0.0000
    Residual |  44.0789788     173  .254791785           R-squared     =  0.3182
-------------+------------------------------           Adj R-squared =  0.3063
       Total |  64.6462215     176  .367308077           Root MSE      =  .50477


-------------------------------------------------------------------------------
     lsalary |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
      lsales |   .1628544   .0392421      4.15   0.000     .0853995    .2403094
     lmktval |    .109243   .0495947      2.20   0.029     .0113545    .2071315
```

```
       ceoten |  .0117054   .0053261    2.20   0.029     .001193    .0222178
        _cons |  4.503795   .2572344   17.51   0.000    3.996073    5.011517
------------------------------------------------------------------------------
```

predict lsal, xb

gen mhat=exp(lsal)

Predicted salary

display _b[_cons]+_b[lsales]*log(5000)+_b[lmktval]*log(10000)+_b[ceoten]*10
7.014077

reg salary mhat, noconstant

```
      Source |       SS       df       MS              Number of obs =     177
-------------+------------------------------           F(  1,   176) =  562.39
       Model |  147352712     1   147352712           Prob > F      =  0.0000
    Residual |  46113900.4   176   262010.798          R-squared     =  0.7616
-------------+------------------------------           Adj R-squared =  0.7603
       Total |  193466612   177   1093031.71           Root MSE      =  511.87

------------------------------------------------------------------------------
      salary |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        mhat |   1.116857   .0470953    23.71   0.000    1.023912    1.209801
------------------------------------------------------------------------------
```

Predicted salary

display _b[mhat]*exp(7.013)
1240.9674

Example 6.8: Predicting CEO Salaries

use http://fmwww.bc.edu/ec-p/data/wooldridge/CEOSAL2

reg salary sales mktval ceoten

```
      Source |       SS       df       MS              Number of obs =     177
-------------+------------------------------           F(  3,   173) =   14.53
       Model |  12230632.6     3   4076877.52          Prob > F      =  0.0000
    Residual |  48535332.2   173   280551.053          R-squared     =  0.2013
-------------+------------------------------           Adj R-squared =  0.1874
```

```
     Total |  60765964.7   176  345261.163            Root MSE      =  529.67

-------------------------------------------------------------------------------
    salary |     Coef.    Std. Err.       t     P>|t|    [95% Conf. Interval]
-----------+-------------------------------------------------------------------
     sales |   .0190191    .0100561     1.89    0.060   -.0008294    .0388676
    mktval |   .0234003    .0094826     2.47    0.015    .0046839    .0421167
    ceoten |   12.70337    5.618052     2.26    0.025    1.614616    23.79211
     _cons |   613.4361    65.23685     9.40    0.000    484.6735    742.1987
-------------------------------------------------------------------------------
```

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**