Stata Textbook Examples
Introductory Econometrics: A Modern Approach by Jeffrey M. Wooldridge (1st & 2nd eds.)
Chapter 9 - More on Specification and Data Problems

Example 9.1: Economic Model of Crime

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

reg narr86 pcnv avgsen tottime ptime86 qemp86 inc86 black hispan

```
      Source |       SS       df       MS              Number of obs =    2725
-------------+------------------------------           F(  8,  2716) =   26.47
       Model |  145.390104     8   18.173763           Prob > F      =  0.0000
    Residual |  1864.95705  2716  .686655763           R-squared     =  0.0723
-------------+------------------------------           Adj R-squared =  0.0696
       Total |  2010.34716  2724  .738012906           Root MSE      =  .82865

------------------------------------------------------------------------------
      narr86 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        pcnv |  -.1332344   .0403502    -3.30   0.001    -.2123546   -.0541141
      avgsen |  -.0113177   .0122401    -0.92   0.355    -.0353185    .0126831
     tottime |   .0120224   .0094352     1.27   0.203    -.0064785    .0305233
     ptime86 |  -.0408417    .008812    -4.63   0.000    -.0581206   -.0235627
      qemp86 |  -.0505398   .0144397    -3.50   0.000    -.0788538   -.0222258
       inc86 |  -.0014887   .0003406    -4.37   0.000    -.0021566   -.0008207
       black |   .3265035   .0454156     7.19   0.000     .2374508    .4155561
      hispan |   .1939144   .0397113     4.88   0.000     .1160469    .2717818
       _cons |   .5686855   .0360461    15.78   0.000     .4980048    .6393661
------------------------------------------------------------------------------
```

reg narr86 pcnv pcnvsq avgsen tottime ptime86 pt86sq qemp86 inc86 inc86sq black hispan

```
      Source |       SS       df       MS              Number of obs =    2725
-------------+------------------------------           F( 11,  2713) =   28.46
       Model |  207.979007    11  18.9071825           Prob > F      =  0.0000
    Residual |  1802.36815  2713   .66434506           R-squared     =  0.1035
-------------+------------------------------           Adj R-squared =  0.0998
       Total |  2010.34716  2724  .738012906           Root MSE      =  .81507

------------------------------------------------------------------------------
      narr86 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        pcnv |   .5525236   .1542372     3.58   0.000     .2500892    .8549579
      pcnvsq |  -.7302119   .1561177    -4.68   0.000    -1.036333   -.4240903
      avgsen |  -.0170216   .0120539    -1.41   0.158    -.0406574    .0066142
     tottime |    .011954   .0092825     1.29   0.198    -.0062474    .0301554
     ptime86 |   .2874334   .0442582     6.49   0.000     .2006501    .3742166
```

```
      pt86sq |  -.0296076   .0038634    -7.66   0.000     -.037183   -.0220321
      qemp86 |  -.0140941   .0173612    -0.81   0.417    -.0481366    .0199485
       inc86 |  -.0034152   .0008037    -4.25   0.000    -.0049912   -.0018392
     inc86sq |   7.19e-06   2.56e-06     2.81   0.005     2.17e-06    .0000122
       black |    .292296     .04483     6.52   0.000     .2043916    .3802004
      hispan |   .1636175   .0394507     4.15   0.000     .0862609     .240974
       _cons |   .5046065   .0368353    13.70   0.000     .4323784    .5768347
------------------------------------------------------------------------------
```

Example 9.2: Housing Price Equation

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

reg price lotsize sqrft bdrms

      Source |       SS       df       MS              Number of obs =      88
-------------+------------------------------           F(  3,    84) =   57.46
       Model |  617130.701     3  205710.234            Prob > F      =  0.0000
    Residual |  300723.805    84   3580.0453           R-squared     =  0.6724
-------------+------------------------------           Adj R-squared =  0.6607
       Total |  917854.506    87  10550.0518           Root MSE      =  59.833

------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lotsize |   .0020677   .0006421     3.22   0.002     .0007908    .0033446
       sqrft |   .1227782   .0132374     9.28   0.000     .0964541    .1491022
       bdrms |   13.85252   9.010145     1.54   0.128     -4.06514    31.77018
       _cons |  -21.77031   29.47504    -0.74   0.462    -80.38466    36.84404
------------------------------------------------------------------------------

predict double r1

gen double r2=r1*r1

gen double r3=r2*r1

reg price lotsize sqrft bdrms r2 r3

      Source |       SS       df       MS              Number of obs =      88
-------------+------------------------------           F(  5,    82) =   39.35
       Model |  647870.698     5   129574.14            Prob > F      =  0.0000
    Residual |  269983.807    82  3292.48546           R-squared     =  0.7059
-------------+------------------------------           Adj R-squared =  0.6879
       Total |  917854.506    87  10550.0518           Root MSE      =   57.38
```

```
------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lotsize |   .0001537    .005203     0.03   0.977    -.0101968    .0105042
       sqrft |   .0175989   .2992508     0.06   0.953    -.5777064    .6129041
       bdrms |   2.174905   33.88811     0.06   0.949    -65.23934    69.58915
          r2 |   .0003534   .0070989     0.05   0.960    -.0137686    .0144755
          r3 |   1.55e-06   6.55e-06     0.24   0.814    -.0000115    .0000146
       _cons |   166.0973   317.4325     0.52   0.602    -465.3772    797.5717
------------------------------------------------------------------------------
```

test r2 r3

```
 ( 1)  r2 = 0.0
 ( 2)  r3 = 0.0

       F(  2,    82) =    4.67
            Prob > F =    0.0120
```

reg lprice llotsize lsqrft bdrms

```
      Source |       SS       df       MS              Number of obs =      88
-------------+------------------------------           F(  3,    84) =   50.42
       Model |  5.15504028     3  1.71834676           Prob > F      =  0.0000
    Residual |  2.86256324    84  .034078134           R-squared     =  0.6430
-------------+------------------------------           Adj R-squared =  0.6302
       Total |  8.01760352    87  .092156362           Root MSE      =   .1846

------------------------------------------------------------------------------
      lprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    llotsize |   .1679667   .0382812     4.39   0.000     .0918404     .244093
      lsqrft |   .7002324   .0928652     7.54   0.000     .5155597    .8849051
       bdrms |   .0369584   .0275313     1.34   0.183    -.0177906    .0917074
       _cons |  -1.297042   .6512836    -1.99   0.050    -2.592191   -.0018931
------------------------------------------------------------------------------
```

predict lphat

gen lph2=lphat*lphat

gen lph3=lphat*lph2

reg lprice llotsize lsqrft bdrms lph2 lph3

```
      Source |       SS       df       MS              Number of obs =      88
-------------+------------------------------           F(  5,    82) =   32.41
       Model |  5.32360126     5  1.06472025           Prob > F      =  0.0000
```

```
    Residual |  2.69400226    82  .032853686           R-squared     =  0.6640
-------------+------------------------------           Adj R-squared =  0.6435
       Total |  8.01760352    87  .092156362           Root MSE      =  .18126

------------------------------------------------------------------------------
      lprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    llotsize |  -4.191584   12.59578    -0.33   0.740     -29.2486    20.86543
      lsqrft |  -17.39336   52.49227    -0.33   0.741    -121.8172     87.0305
       bdrms |  -.9276645    2.76988    -0.33   0.739    -6.437838    4.582509
        lph2 |   3.921189   13.01484     0.30   0.764    -21.96948    29.81186
        lph3 |  -.1933951   .7521095    -0.26   0.798     -1.68958     1.30279
       _cons |   88.08799   240.9851     0.37   0.716    -391.3081    567.4841
------------------------------------------------------------------------------
```

test lph2 lph3

```
 ( 1)  lph2 = 0.0
 ( 2)  lph3 = 0.0

       F(  2,    82) =    2.57
            Prob > F =    0.0831
```

Example 9.3: IQ as a Price for Ability

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

gen educIQ=educ*IQ

reg lwage educ exper tenure married south urban black

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  7,   927) =   44.75
       Model |  41.8377677     7  5.97682396           Prob > F      =  0.0000
    Residual |  123.818527   927  .133569069           R-squared     =  0.2526
-------------+------------------------------           Adj R-squared =  0.2469
       Total |  165.656294   934  .177362199           Root MSE      =  .36547

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0654307   .0062504    10.47   0.000     .0531642    .0776973
       exper |    .014043   .0031852     4.41   0.000      .007792     .020294
      tenure |   .0117473    .002453     4.79   0.000     .0069333    .0165613
     married |   .1994171   .0390502     5.11   0.000     .1227802    .2760541
       south |  -.0909036   .0262485    -3.46   0.001     -.142417   -.0393903
```

```
       urban |   .1839121   .0269583     6.82   0.000     .1310056    .2368185
       black |  -.1883499   .0376666    -5.00   0.000    -.2622717   -.1144282
       _cons |   5.395497    .113225    47.65   0.000      5.17329    5.617704
------------------------------------------------------------------------------
```

reg lwage educ exper tenure married south urban black IQ

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  8,   926) =   41.27
       Model |  43.5360229     8  5.44200287           Prob > F      =  0.0000
    Residual |  122.120271   926  .131879343           R-squared     =  0.2628
-------------+------------------------------           Adj R-squared =  0.2564
       Total |  165.656294   934  .177362199           Root MSE      =  .36315

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0544106   .0069285     7.85   0.000     .0408133     .068008
       exper |   .0141458   .0031651     4.47   0.000     .0079342    .0203575
      tenure |   .0113951   .0024394     4.67   0.000     .0066077    .0161825
     married |   .1997644   .0388025     5.15   0.000     .1236134    .2759154
       south |  -.0801695   .0262529    -3.05   0.002    -.1316916   -.0286473
       urban |   .1819463   .0267929     6.79   0.000     .1293645    .2345281
       black |  -.1431253   .0394925    -3.62   0.000    -.2206304   -.0656202
          IQ |   .0035591   .0009918     3.59   0.000     .0016127    .0055056
       _cons |   5.176439   .1280006    40.44   0.000     4.925234    5.427644
------------------------------------------------------------------------------
```

reg lwage educ exper tenure married south urban black IQ educIQ

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  9,   925) =   36.76
       Model |  43.6401304     9  4.84890337           Prob > F      =  0.0000
    Residual |  122.016164   925  .131909366           R-squared     =  0.2634
-------------+------------------------------           Adj R-squared =  0.2563
       Total |  165.656294   934  .177362199           Root MSE      =  .36319

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0184558   .0410608     0.45   0.653    -.0621273     .099039
       exper |   .0139072   .0031768     4.38   0.000     .0076725    .0201418
      tenure |   .0113929   .0024397     4.67   0.000     .0066049    .0161808
     married |   .2008658   .0388267     5.17   0.000     .1246672    .2770644
       south |  -.0802354    .026256    -3.06   0.002    -.1317637   -.0287071
       urban |   .1835758   .0268586     6.83   0.000     .1308649    .2362867
       black |  -.1466989   .0397013    -3.70   0.000    -.2246139   -.0687839
          IQ |  -.0009418   .0051625    -0.18   0.855    -.0110734    .0091899
       educIQ |   .0003399   .0003826     0.89   0.375    -.0004109    .0010907
```

```
       _cons |   5.648249   .5462963    10.34   0.000     4.576125    6.720373
------------------------------------------------------------------------------
```

Example 9.4: City Crime Rates

use http://fmwww.bc.edu/ec-p/data/wooldridge/CRIME2

reg lcrmrte unem llawexpc if d87==1

```
      Source |       SS       df       MS              Number of obs =      46
-------------+------------------------------           F(  2,    43) =    1.30
       Model |  .271987199     2    .1359936           Prob > F      =  0.2824
    Residual |  4.48998214    43  .104418189           R-squared     =  0.0571
-------------+------------------------------           Adj R-squared =  0.0133
       Total |  4.76196934    45  .105821541           Root MSE      =  .32314

------------------------------------------------------------------------------
     lcrmrte |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        unem |  -.0290032   .0323387    -0.90   0.375    -.0942205    .0362141
    llawexpc |   .2033652   .1726534     1.18   0.245    -.1448236    .5515539
       _cons |   3.342899   1.250527     2.67   0.011     .8209721    5.864826
------------------------------------------------------------------------------
```

reg lcrmrte unem llawexpc lcrmrt_1

```
      Source |       SS       df       MS              Number of obs =      46
-------------+------------------------------           F(  3,    42) =   29.73
       Model |  3.23732846     3  1.07910949           Prob > F      =  0.0000
    Residual |  1.52464088    42  .036300973           R-squared     =  0.6798
-------------+------------------------------           Adj R-squared =  0.6570
       Total |  4.76196934    45  .105821541           Root MSE      =  .19053

------------------------------------------------------------------------------
     lcrmrte |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        unem |    .008621   .0195166     0.44   0.661    -.0307652    .0480072
    llawexpc |  -.1395764   .1086412    -1.28   0.206    -.3588231    .0796704
    lcrmrt_1 |   1.193923   .1320985     9.04   0.000     .9273371    1.460508
       _cons |   .0764511   .8211433     0.09   0.926    -1.580683    1.733585
------------------------------------------------------------------------------
```

Example 9.5: Saving Function with Measurement Error

Dataset is not provided

Example 9.6: Measurement Error in Scrap Rates

Dataset is not provided

Example 9.7: GPA Equation with Measurement Error

Dataset is not provided

Example 9.8: R&D Intensity and Firm Size

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

reg rdintens sales profmarg

```
      Source |       SS       df       MS              Number of obs =      32
-------------+------------------------------           F(  2,    29) =    1.19
       Model |  8.28423732     2  4.14211866           Prob > F      =  0.3173
    Residual |  100.549233    29  3.46721493           R-squared     =  0.0761
-------------+------------------------------           Adj R-squared =  0.0124
       Total |   108.83347    31  3.51075711           Root MSE      =   1.862

------------------------------------------------------------------------------
    rdintens |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       sales |   .0000534   .0000441     1.21   0.236    -.0000368    .0001435
    profmarg |   .0446166   .0461805     0.97   0.342    -.0498332    .1390664
       _cons |   2.625261   .5855328     4.48   0.000     1.427712     3.82281
------------------------------------------------------------------------------
```

reg rdintens sales profmarg if sales<20000

```
      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  2,    28) =    2.92
       Model |  18.7880289     2  9.39401445           Prob > F      =  0.0702
    Residual |  89.9330615    28  3.21189505           R-squared     =  0.1728
-------------+------------------------------           Adj R-squared =  0.1137
       Total |   108.72109    30  3.62403635           Root MSE      =  1.7922

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

```
-------------+----------------------------------------------------------------
       sales |   .0001856   .0000842     2.20   0.036     .0000131    .0003581
    profmarg |   .0478411   .0444831     1.08   0.291    -.0432784    .1389605
       _cons |   2.296851   .5918045     3.88   0.001     1.084594    3.509107
------------------------------------------------------------------------------
```

Example 9.9: R&D Intensity

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

reg lrd lsales profmarg

```
      Source |       SS       df       MS              Number of obs =      32
-------------+------------------------------           F(  2,    29) =  162.24
       Model |   85.597056     2   42.798528           Prob > F      =  0.0000
    Residual |   7.6502049    29  .263800169           R-squared     =  0.9180
-------------+------------------------------           Adj R-squared =  0.9123
       Total |  93.2472609    31  3.00797616           Root MSE      =  .51361

------------------------------------------------------------------------------
         lrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsales |   1.084228   .0601941    18.01   0.000     .9611173    1.207339
    profmarg |   .0216594    .012782     1.69   0.101    -.0044827    .0478015
       _cons |  -4.378349   .4680132    -9.36   0.000    -5.335544   -3.421155
------------------------------------------------------------------------------
```

reg lrd lsales profmarg if sales<20000

```
      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  2,    28) =  131.42
       Model |  71.7655416     2  35.8827708           Prob > F      =  0.0000
    Residual |  7.64489638    28  .273032014           R-squared     =  0.9037
-------------+------------------------------           Adj R-squared =  0.8969
       Total |   79.410438    30   2.6470146           Root MSE      =  .52252

------------------------------------------------------------------------------
         lrd |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsales |   1.088057   .0671128    16.21   0.000     .9505826    1.225531
    profmarg |    .021759   .0130233     1.67   0.106     -.004918     .048436
       _cons |  -4.404225   .5110168    -8.62   0.000    -5.450995   -3.357454
------------------------------------------------------------------------------
```

Example 9.10: State Infant Mortality Rates

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

reg infmort lpcinc lphysic lpopul if year==1990

      Source |       SS       df       MS              Number of obs =      51
-------------+------------------------------           F(  3,    47) =    2.53
       Model |  32.1624527     3  10.7208176           Prob > F      =  0.0684
    Residual |  199.085016    47  4.23585141           R-squared     =  0.1391
-------------+------------------------------           Adj R-squared =  0.0841
       Total |  231.247469    50  4.62494938           Root MSE      =  2.0581

------------------------------------------------------------------------------
     infmort |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lpcinc |  -4.684585   2.604134    -1.80   0.078    -9.923426    .5542562
     lphysic |   4.153227   1.512663     2.75   0.009     1.110143    7.196312
      lpopul |  -.0878245   .2872503    -0.31   0.761    -.6656976    .4900486
       _cons |   33.85875   20.42792     1.66   0.104    -7.236927    74.95444
------------------------------------------------------------------------------

reg infmort lpcinc lphysic lpopul if infmort<20 & year==1990

      Source |       SS       df       MS              Number of obs =      50
-------------+------------------------------           F(  3,    46) =    5.76
       Model |  26.8600392     3  8.95334639           Prob > F      =  0.0020
    Residual |  71.4631627    46  1.55354702           R-squared     =  0.2732
-------------+------------------------------           Adj R-squared =  0.2258
       Total |  98.3232019    49  2.00659596           Root MSE      =  1.2464

------------------------------------------------------------------------------
     infmort |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lpcinc |  -.5669247   1.641217    -0.35   0.731    -3.870523    2.736674
     lphysic |   -2.74184   1.190771    -2.30   0.026    -5.138737    -.344943
      lpopul |   .6292351   .1911062     3.29   0.002     .2445584    1.013912
       _cons |   23.95478   12.41949     1.93   0.060    -1.044345    48.95391
------------------------------------------------------------------------------
```

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**