GOVT 613—Schaffner
Testing for Serial Correlation and Plotting Residuals in Stata

One of the important things that we will be focusing on this week and next is the examination of your residuals—the difference between your actual values of Y and your predicted values of Y. Learning to plot your residuals will allow you to look for problems such as serial correlation and heteroskedasticity. This handout deals with serial correlation, but you will see how you can use some of these methods to look for heteroskedasticity next week.

Durbin-Watson Test

The Durbin-Watson Test allows you to examine whether first order serial correlation is a problem in your regression model (see Studenmund pp. 325-329). To follow along with this example, access the time series dataset at http://nw08.american.edu/~schaffne/trust2.dta. *Trust2* is a variable measuring the percentage of the American public that reported trusting the federal government during a given quarter. *Prezapprv* measures the percentage approving of the president, *sentiment* is the consumer sentiment index for that quarter, and *mood* is a measure of whether the public wants more or less government during that quarter. The variable *quarter* is our measure of time.

```
. use trust2

. reg trust2 prezapprv sentiment mood

      Source |       SS       df       MS              Number of obs =      63
-------------+------------------------------           F(  3,    59) =   20.56
       Model |   2678.3999     3  892.799966           Prob > F      =  0.0000
    Residual |  2561.99981    59  43.4237256           R-squared     =  0.5111
-------------+------------------------------           Adj R-squared =  0.4862
       Total |  5240.39971    62   84.522576           Root MSE      =  6.5897

------------------------------------------------------------------------------
      trust2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   prezapprv |   .2470662   .1055555     2.34   0.023     .0358502    .4582823
   sentiment |   .3271492   .1056295     3.10   0.003      .115785    .5385134
        mood |    1.89474   .3422136     5.54   0.000     1.209972    2.579508
       _cons |  -123.3015    24.2443    -5.09   0.000    -171.8143   -74.78878
------------------------------------------------------------------------------
```

We expect first order serial correlation here since the level of trust in a previous quarter is probably related to trust in the next quarter. To test for this, we ask Stata to give us the Durbin-Watson Statistic.

```
. dwstat

Durbin-Watson d-statistic(  4,    63) =  .7168112
```

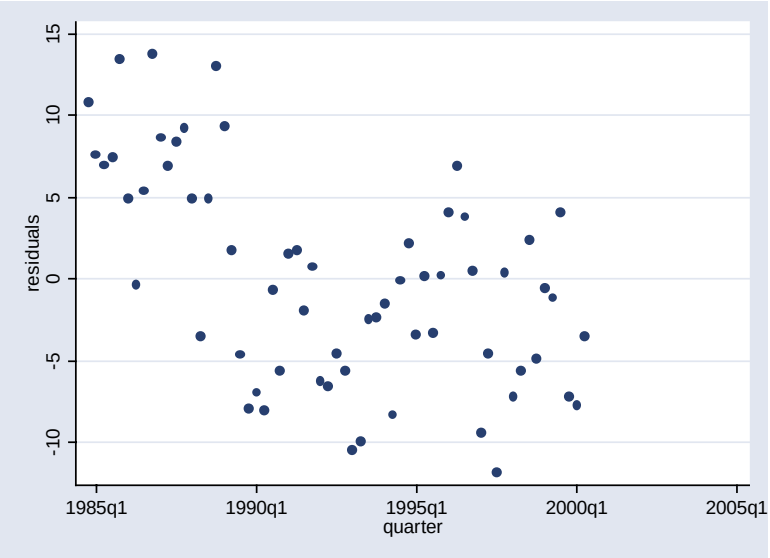
Recall that we are looking for a statistic around 2 if we have no autocorrelation. This statistic is far from 2 and indicates positive first order autocorrelation. If you check the Durbin-Watson table in the back of your book, you'll see that the critical statistic for 95% confidence is 1.40. Our value of .7168 lies far beyond that indicating positive first order serial correlation. In other words, a previous value of the error term is likely to be positively associated with the current value.

Plotting the Residuals

Now let's take a look at our residuals to get a better sense of how the residuals are related to each other over time. To do this, we need to generate a variable that equals the residuals for our model. We do this with the predict command.

```
. predict residuals, resid
(9 missing values generated)

. scatter residuals quarter
```



This scatter plot shows clearly that previous values of the error term are associated with the current value of that term. One way to deal with this is the use of Newey-West standard errors (Studenmund pp. 334-5). To generate output with these standard errors use the *newey* command. Note that you need to specify the lag (which is 1 here because you have first order autocorrelation).

```
. newey trust2 prezapprv sentiment mood, lag(1)

Regression with Newey-West standard errors          Number of obs  =        63
maximum lag: 1                                      F(  3,    59)  =     21.66
                                                    Prob > F       =    0.0000

------------------------------------------------------------------------------
             |             Newey-West
      trust2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   prezapprv |   .2470662    .112748     2.19   0.032     .0214581    .4726744
   sentiment |   .3271492   .0986392     3.32   0.002     .1297726    .5245259
        mood |    1.89474   .3921366     4.83   0.000     1.110076    2.679404
       _cons |  -123.3015    25.6912    -4.80   0.000    -174.7095   -71.89354
------------------------------------------------------------------------------
```

Notice that the standard errors for *prezapprv* and *mood* are now larger, although, the standard error for sentiment actually decreased. Also, only the standard errors change, not the coefficients, because only the standard errors were biased by the serial correlation.