DEPARTMENT OF POLITICAL SCIENCE
AND
INTERNATIONAL RELATIONS
Posc/Uapp 816

TIME SERIES REGRESSION

I. AGENDA:
- A. A couple of general considerations in analyzing time series data
- B. Intervention analysis
  1. Example of simple interupted time series models.
  2. Checking the adequacy of the models.
  3. Modification.

II. SOME PROBLEMS IN ANALYZING TIME SERIES:
- A. In the last class (Class 19) we used regression to see how an "intervention" affected a dependent variable measured at discrete time periods.
  1. We'll continue that analysis in a moment.
  2. But we first need to review the assumptions underlying regression analysis, particularly those pertaining to the error term.
- B. Regression assumptions:
  1. If "time" is the unit of analysis we can still regress some dependent variable, Y, on one or more independent variables.
     i. Last time we dealt with a particularly simple variable, a "time counter."
        1) That is, X was defined as $X_t = 1, 2, 3, ..., N$.
     ii. The form of a regression model with one explanatory variable is:

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

  2. Assumptions about ε_t, the "error term":
     i. $E(\varepsilon_t) = 0$, zero mean
     ii. $E(\varepsilon_t^2) = \sigma^2$, constant variance
     iii. $E(\varepsilon_t, X_t) = 0$, no correlation with X
     iv. $E(\varepsilon_t, \varepsilon_{t-1})$, no autocorrelation.
     v. ε_t ~ Normally distributed (for hypothesis testing).
  3. Assumption four is especially important and most likely not to be met when using time series data.
- C. Autocorrelation.
  1. It is not uncommon for errors to "track' themselves; that is, for the error a time t to depend in part on its value at t - m, where m is a prior time period.

i. The most common situation occurs when m = 1, which is called a first-order autocorrelation

$$\varepsilon_t \varepsilon_{t-1} \neq 0$$

1) This form indicates that the errors at a prior time (i.e., one prior time period) are correlated (not independent) of errors at the following time period.

2. Another view of autocorrelation.

i. Suppose that the errors in one time period are correlated with the errors in the preceding time period.

1) This is actually a common occurrence.

ii. Such a situation can be called a **first-order autoregressive** process:

1) A simple linear model has the usual form

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

2) But now the errors are related by the (linear) simple regression function:

$$\varepsilon_t = \rho \varepsilon_{t-1} + \nu_t$$

3) That is the error at time t is a function of the error at time t - 1 and a **random** disturbance.

iii. This model has these properties. (That is, we make these assumptions.)

1) $E[\nu_t] = 0$; the mean of disturbances is zero ("they cancel out");
2) $E[\nu_t^2] = \sigma_\nu^2$; the disturbances have constant variance;
3) $E[\nu_t, \nu_{t-1}] = 0$, disturbances are uncorrelated;
4) $E[\nu_t, \varepsilon_{t-1}] = 0$; no correlation between disturbance at t and the error in the model at t - 1.
5) $-1 < \rho < +1$; rho, the autocorrelation parameter is a fraction.

iv. Since each error, ε_t, is a portion of the immediately preceding one plus a random disturbance, it can be written as (using repeated substitutions):

$$
\begin{aligned}
\varepsilon_t &= \rho\varepsilon_{t-1} + \nu_t \\
&= \rho[\rho\varepsilon_{t-2} + \nu_{t-1}] + \nu_t \\
&= \rho^2\varepsilon_{t-2} + \rho\nu_{t-1} + \nu_t \\
&= \rho^2[\rho\varepsilon_{t-3} + \nu_{t-2}] + \rho\nu_{t-1} + \nu_t \\
&= \rho^3\varepsilon_{t-3} + \rho^2\nu_{t-1} + \rho\nu_{t-1} + \nu_t \\
&\textit{etc.} \\
&= \rho^t\varepsilon_0 + \rho^{t-1}\nu_{t-m+1} + \ldots + \rho\nu_{t-1} + \nu_0
\end{aligned}
$$

3. In other words, the errors at time t, are created as a linear function of a random disturbance and ultimately of the "first" or original error, ε_0
4. If we knew the value of the autocorrelation parameter, ρ, we would be in a position to specify the error structure and use OLS to estimate the parameters of the time series regression model.
   i. But since we don't a problem arises.

D. The consequences of autocorrelation.
1. Recall that an estimator is **unbiased** if its expected value equals the population parameter it is estimating.
   i. Example: the mean is an unbiased estimator of the population mean because $\boldsymbol{E(\bar{Y}) = \mu}$
2. But of course estimators have variances; that is, they vary from sample to sample, a fact represented by the standard error of the estimator.
3. (Positive) autocorrelation has the effect of deflating the size of standard errors. This decrease in turn means that observed t values will be too large, leading one to reject null hypotheses that should perhaps be accepted, and that confidence intervals will be too narrow.
4. How serious this problem is depends on what one is doing with the data.
   i. For a large sample (T is large) and only one estimation equation the potentially misleading results may be outweighed by the simplicity of ignoring the problem
   ii. Usually, however, social scientists and policy analysts build and test lots of models with the same data and so conduct numerous tests. In this situation, drawing firm conclusions might be difficult, especially in view of all the other problems inherent in statistical analysis.

5. So later we'll look at method for determining whether autocorrelation seems to be a problem and what to do it is. But for now back to some simple intervention analysis.

III. EXAMPLE INTERVENTION MODEL: ENERGY IMPORTS AND EXPORTS

A. Let's examine some data from the Department of Energy regarding petroleum products imports over the (nearly) last half century.

1. In particular we may want to know if the oil crises of the early 1970s had any effect on imports.
2. Here are the data. The dependent variable is petroleum imports in millions of barrels a day.

Year	Imports	Year	Imports
1949	0.22	1973	3.01
1950	0.36	1974	2.64
1951	0.35	1975	1.95
1952	0.38	1976	2.03
1953	0.39	1977	2.19
1954	0.40	1978	2.01
1955	0.47	1979	1.94
1956	0.50	1980	1.65
1957	0.55	1981	1.60
1958	0.75	1982	1.63
1959	0.81	1983	1.72
1960	0.80	1984	2.01
1961	0.87	1985	1.87
1962	0.96	1986	2.05
1963	0.99	1987	2.00
1964	1.06	1988	2.30
1965	1.23	1989	2.22
1966	1.35	1990	2.12
1967	1.41	1991	1.84
1968	1.55	1992	1.80
1969	1.76	1993	1.83
1970	2.10	1994	1.93
1971	2.25	1995	1.61
1972	2.53	1996	1.92

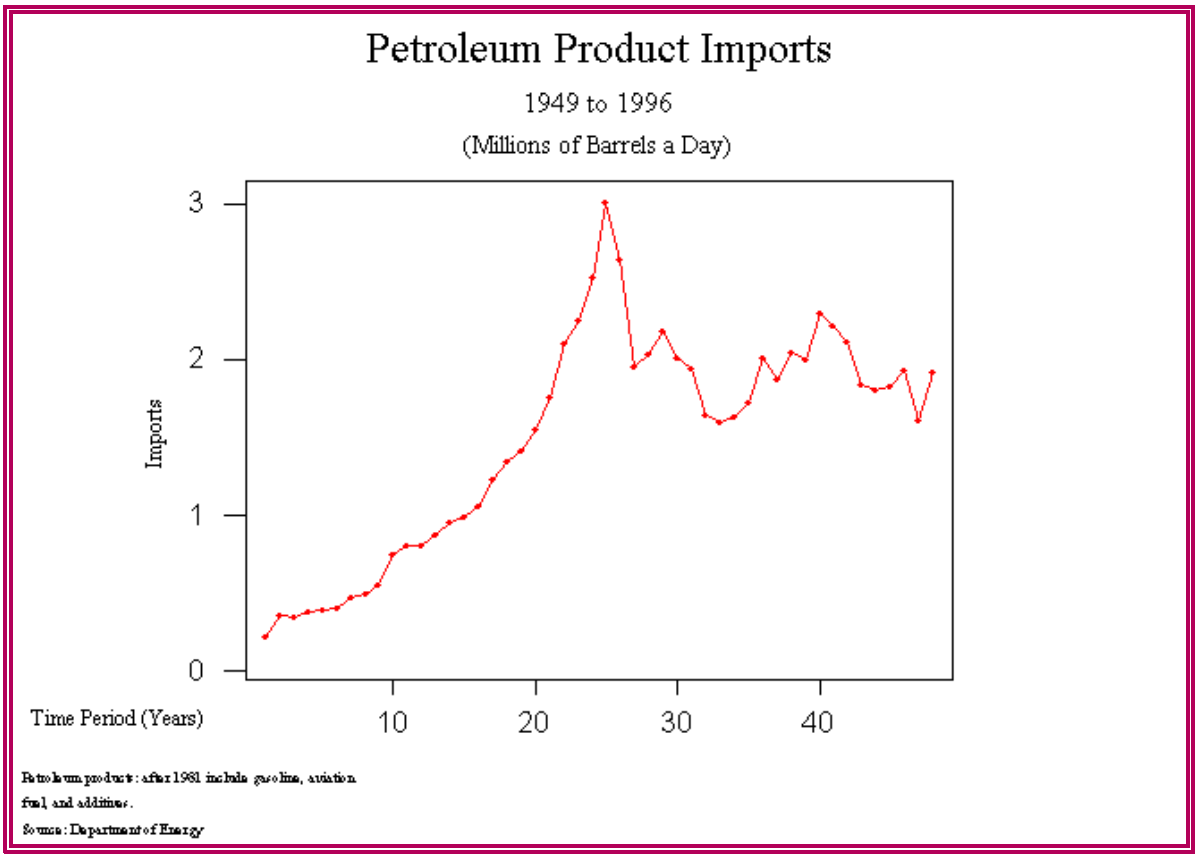
Table 1: Petroleum Product Imports

i. We can start by graphing the data as usual.

1) For these it is convenient to use time series plot, although the data could be plotted as Y versus year.

a) Doing so, in fact, might help us interpret the data.

b) The next page contains a time series plot.



2) The plot shows an obvious point: petroleum imports leveled off after about 1972 to 1974.
   a) We'll use the year 1971, the beginning of the first postwar American "energy" crisis.
   b) What impact did it have on imports?

3. We can compare the before and after levels using usual methods such as difference of means procedures and tests or analysis of variance.
   i. Recall that analysis of variance allows one to compare and test for differences of two or more means.
   ii. We looked at the procedure when examining dummy variables.
   iii. Below are the results using MINITAB.

```
One-way Analysis of Variance

Analysis of Variance for Petroimp
Source     DF        SS        MS        F        P
Petdumy     1     9.034     9.034    27.73    0.000
Error      46    14.984     0.326
Total      47    24.019
                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
Level       N      Mean     StDev
---------+---------+---------+-------
0          25    1.0820    0.7544  (-----*-----)
1          23    1.9504    0.2453                      (-----*-----)

---------+---------+---------+-------
Pooled Standard Dev =   0.5707          1.20        1.60     2.00
```

Table 2: ANOVA for Petroleum Data

iv. Clearly the U.S. imported more oil in the post-crisis period than before.

B. This fact can be verified by the simple change in level model discussed last time (Class 19).

1. Let $X_1 = 0$ for time periods before 1971 and 1 afterwards.
2. The results of the regression of imports on X_1 alone are:

```
The regression equation is
Petroimp = 1.08 + 0.868 Petdumy

Predictor        Coef       StDev          T        P
Constant       1.0820      0.1141       9.48    0.000
Petdumy        0.8684      0.1649       5.27    0.000

S = 0.5707      R-Sq = 37.6%     R-Sq(adj) = 36.3%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      9.0344      9.0344     27.73    0.000
Residual Error    46     14.9845      0.3257
Total             47     24.0189
```

Table 3: Regression for Petroleum Data

i. To interpret the data, which contains the same information as Table 2, substitute 0 for X_1 and observe that the expected level of imports prior to the crisis is 1.083 million barrels a day.

ii. After 1971 when $X_1 = 1$ the expected level is 1.083 + .868 = 1.951 million barrels.

1) The regression parameter (.8684) in this case gives the effect of the intervention on the level of imports.

iii. But what about trends of patterns in importation?

C. To investigate this question let's use a model that includes changes in level and slope or trend.

1. Let $X_1 = 1, 2, 3,...48$, a counter for year or time period;
2. Now let $X_2 = 0$ for periods before 1971 and 1 otherwise;
3. and let $X_3 = X_1 X_2$, an "interaction" variable that creates a dummy counter of 0 before the intervention and time period number after.
4. The model is:

$$E(Y)_t = \beta_0 + \beta_1 X_t + \beta_2 X_2 + \beta_3 X_3$$

i. Here β_0 is the initial level of imports; β_1 is the trend in the initial (pre-intervention) period; and β_2 and β_3 are the effects of the intervention on the levels and trend.

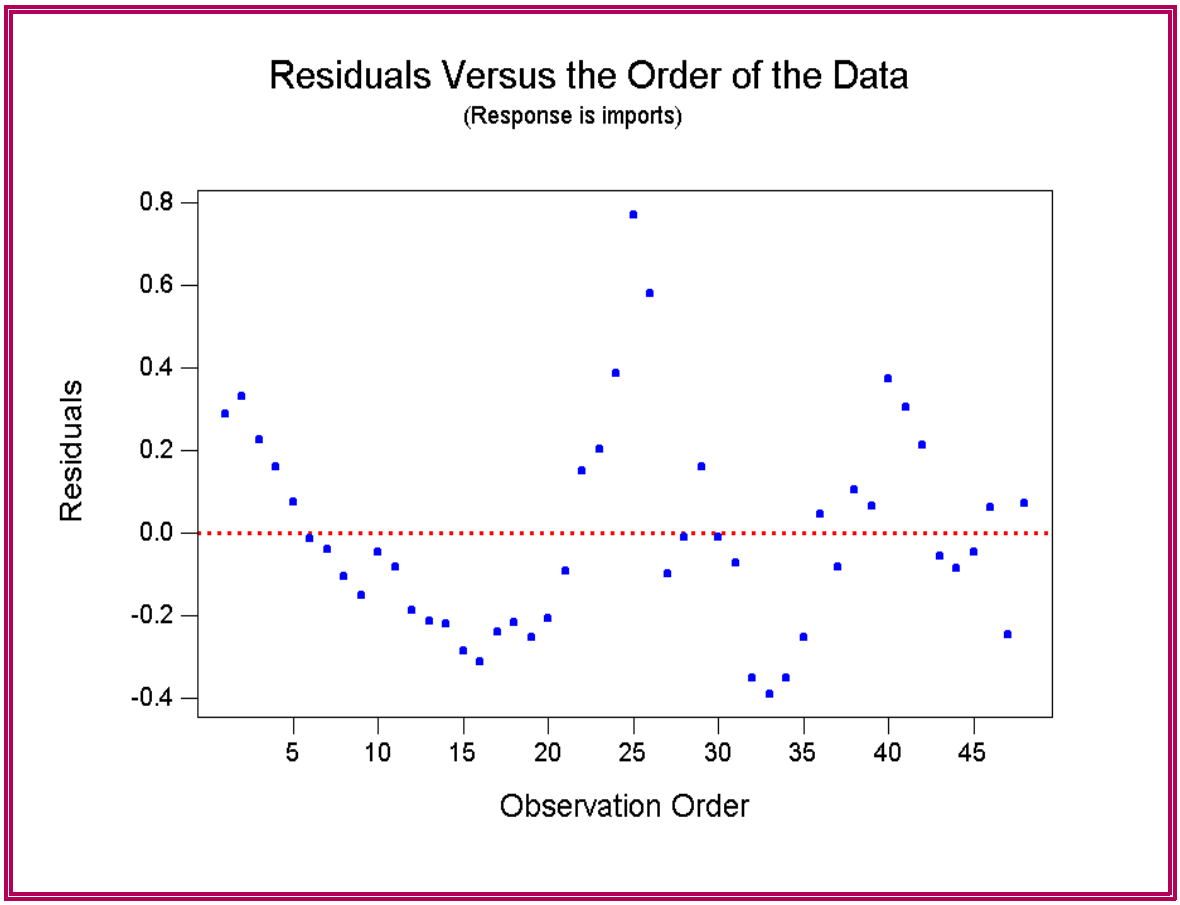
5. The results of the regression analysis are:

```
The regression equation is
Petroimp = - 0.167 + 0.0961 Counter + 2.47 Petdumy - 0.106 Petinter

Predictor        Coef       StDev          T        P
Constant      -0.1670      0.1057      -1.58    0.121
Counter      0.096077    0.007109      13.51    0.000
Petdumy        2.4732      0.3208       7.71    0.000
Petinter     -0.10569     0.01075      -9.84    0.000

S = 0.2563      R-Sq = 88.0%     R-Sq(adj) = 87.1%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         3     21.1280      7.0427    107.19    0.000
Residual Error    44      2.8909      0.0657
Total             47     24.0189


Durbin-Watson statistic = 0.53
```

Table 4: Multiple Regression for Petroleum Data

6. At very first glance the model seems to fit the data and makes sense given our expectations and the time series plot.
   i. Note in particular the slope or trend.
      1) In the pre-crisis period the slope is +.096 million barrels a day.
      2) In the post period it drops to .096077 - .10569 = -.00961.
      3) There has thus been a change from a strong rate of importation to a much smaller (and decreasing) rate.

D. Model adequacy.
1. But just how adequate is the model.
2. Actually, it makes sense substantively but most social scientists would have trouble accepting it before seeing additional data.
3. In particular look at the errors plot against time order (or the appearance of the data):



i. The residuals, which are estimators of the errors, clearly follow a pattern. In fact the pattern is obvious: positive residuals follow positive residuals and negative follow negative.
ii. This is an indication of positive autocorrelation.

IV. DURBIN WATSON TEST FOR AUTOCORRELATION:

A. Identifying serial correlation: the Durbin-Waston test:

1. The previous regression table contained a new statistic, the Durbin-Waston test.
   i. It is used to test the hypothesis that the autocorrelation parameter, ρ is zero.
   ii. That is,

$$H_0: \rho = 0$$

versus (for positive autocorrelation)

$$H_A: \rho > 0$$

2. The statistic is:

$$DW = \frac{\sum_{t=2}^{N} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{N} \hat{\varepsilon}_t^2}$$

   i. The $\hat{\varepsilon}'s$ are residuals based on an estimated model.

3. This formula can perhaps be interpreted best as a correlation among lagged residuals.

B. Lags:

1. The only tricky term in this formula is the subtraction of a value at time t by a previous value.
2. In the case of first order autocorrelation we need to worry about just one lag.
   i. If the previous value is just one time period before (as in this case), the term is called **lag 1**. Lags of 2 and 3 are occasionally used. The

figure below illustrates the idea of "lagged" variables.

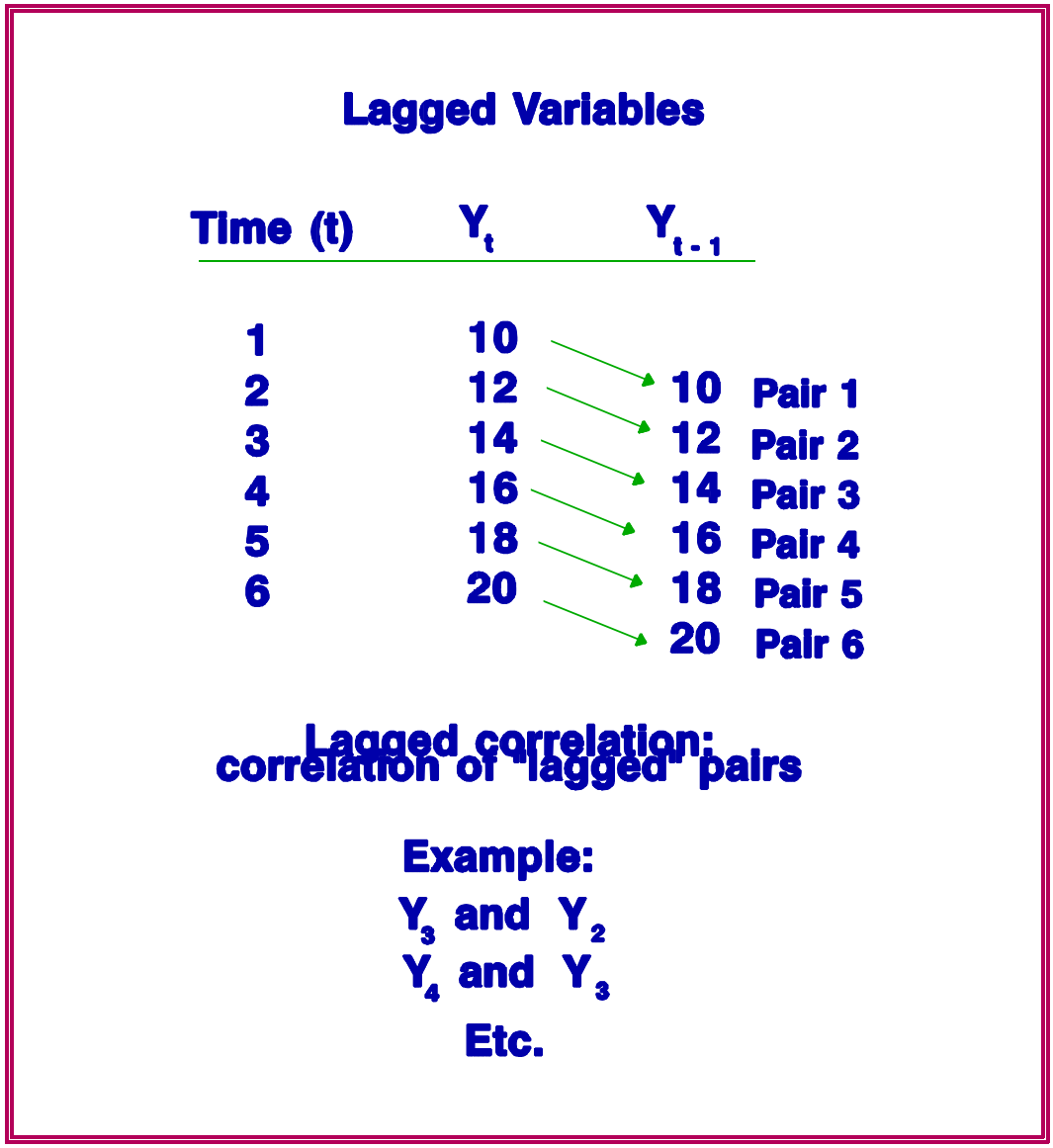


Figure 3: Lagged Variables

C. Essentially, we shift all the data points down one time period (losing one in the process) to create another variable. Hence, the pairs to be correlated are Y at time 1 and Y at 2; Y at time 2 and Y at 3; Y at time 3 and Y at 4, and so forth.

1. These are the values that appear in the formula for the Durbin-Watson statistic.
2. The idea of lagging is also used below to transform Y and X.
3. Note that we lose a case each time we lag.

D. If $\rho = 0$, then the DW statistic will equal 2, except for sampling error. If $\rho = 1$, the DW statistic = 0.

1. Unfortunately, intermediate sample values are not tested in the usual fashion of comparing them with some critical value. Instead, the following rules of thumb:
   i. If ρ = +, then DW equals about 0.
   ii. If ρ = -1, then DW equals about 4.
   iii. If ρ = 0 then DW equals about 2.
      1) A value close to 2, say 1.80, suggests that autocorrelation may not be a problem.
   iv. To evaluate a hypothesis one uses a table of DW values.
      1) Find the N (or T), the number of time periods and K, the number of independent variables.
      2) Use these two numbers to find two values: lower bound **L**, and an upper bound **U**.
   v. If DW is smaller than the lower bound, conclude that the autocorrelation coefficient is positive and reject the null hypothesis.
      1) Transform the data as suggested below.
   vi. If the observed DW statistic is greater than the upper bound, conclude that the autocorrelation is 0.
   vii. If DW lies between **U** and **L**, the test is inconclusive. I would suggest treating the data as though there were serial correlation, although strictly speaking the null hypothesis cannot be reject.
2. What are the upper and lower bounds? Where do we get them.
   i. Most texts on time series analysis and forecasting provide them. I will attempt to scan a copy and make it available. In the meantime, one will be attached to the next set of

E. An alternative:

1. Since the Durbin-Watson test is often inconclusive, some authors suggest simply correlated the estimated errors (i.e., the residuals, $\hat{\varepsilon}_t$) and the lagged residuals, $\hat{\varepsilon}_{t-1}$. If the correlation is greater than .3, assume positive autocorrelation.

F. Notes: this is a very conservative procedure in that we could accept some null hypotheses (that is, have too large intervals) more than we should.

V. DEALING WITH SERIAL CORRELATION:

A. Several procedures have been recommended for dealing with the effects of autocorrelation. The one we use here, called the Cochrane-Orcutt procedure, sometimes involves an "iterative" method: one first estimates a model, then uses the residuals from this model to estimate the autocorrelation parameter, ρ, by lagging the them. Next one adjusts the original data, computes parameter estimates

(the estimated β's). If there is evidence of autocorrelation, one obtains new estimates of the residuals using the revised model and estimates ρ again. The data are once more transformed using this new estimate of the autocorrelation parameter. The process can be repeated until the serial correlation disappears as much as possible (it might never).

1. In fact, however, we usually need only one or two iterations.

B. More specifically, the steps are:

1. Use OLS to obtain the residuals. Store these in somewhere.
2. Lag the residuals 1 time period to obtain $\hat{\varepsilon}_{t-1}$. MINITAB has a procedure:
   i. Go to **Statistics**, **Time series**, then **Lag**.
   ii. Pick the column containing the residuals from the model.
   iii. Pick another column to store the lag residuals and press OK.
   iv. You should name these columns clearly to keep the book keeping simple.
      1) Also, look at the data window to see what's going on.
3. Use descriptive statistics to obtain the simple correlation between the residual column and the lagged residuals.
   i. **This is the estimate of ρ, the autocorrelation parameter.**
4. You will also have to lag the dependent and independent variables so they can be transformed.
   i. Make sure you know where they are stored.
5. Now transform the dependent variable:

$$Y_t^* = Y_t - \hat{\rho} Y_{t-1}$$

   i. That is, create a new variable (with the calculator or mathematical expressions or **let** command) that is simply Y at time t minus ρ times Y at time t - 1.
      1) Example:
         a) Suppose Y at time t is stored in column 10, its lagged version in column 15 and the estimated autocorrelation is .568. Then the let command will store the new Y in column 25:

```
mtb>let c25 = c10 - (.568)*c15
```

6. The first independent variable is transformed by

$$X_t^* = X_t - \hat{\rho} X_{t-1}$$

i. When using MINITAB just follow the procedure describe above for Y.
   1) Example:
      a) If X_1, the counter say, is stored in column 7 and its lag in column 17 (and once again the estimated ρ is .568) then the MINITAB command will be

```
mtb>let c26 = c7 - (.568)*c17
```

ii. Since we are dealing with lagged variables, we will lose one observation (see the figure above).
   1) For now we can forget about them or use the following to replace these missing first values.

$$Y_1^* = \sqrt{1 - \hat{\rho}^2} Y_1$$

$$X_{1,1}^* = \sqrt{1 - \hat{\rho}^2} X_{1,1}$$

$$X_{2,1}^* = \sqrt{1 - \hat{\rho}^2} X_{2,1}$$

iii. The Y_1^* ***and*** $X_{1,1}^*$ mean the <u>first</u> case for Y and X_1 and so forth.

   1) If there are more X's proceed in the same way.

7. Finally, regress the transformed variables (i.e., Y^* on the X^*'s to obtain new estimates of the coefficients, residuals, and so forth.
   i. Check the model adequacy with the Durbin-Watson statistic, plots of errors and the usual.
   ii. If the model is not satisfactory, treat the Y* and X* as "raw" data and go through the process again.
8. Usually, you will have to do this only once.

VI. NEXT TIME:
   A. Example of the iterative procedure
   B. More discussion of time series.

Go to Notes page

Go to Statistics page