Chapter 9: Serial Correlation

In this chapter:

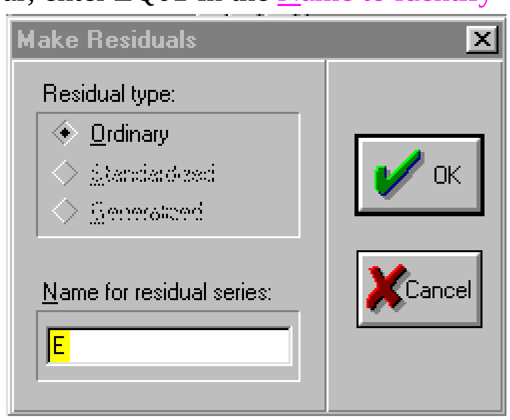
1. Creating a residual series from a regression model
2. Plotting the error term to detect serial correlation (*UE*, pp. 313-315)
3. Using regression to estimate ρ, the first order serial correlation coefficient (*UE*, Equation 9.1, pp. 311-312)
4. Viewing the Durbin-Watson d statistic in the EViews Estimation Output window (*UE* 9.3)
5. Estimating generalized least squares using the AR(1) method (*UE* 9.4.2)
6. Estimating generalized least squares (GLS) equations using the Cochrane-Orcutt method (*UE* 9.4.2)
7. Exercise

Serial correlation analysis involves an examination of the error term. The demand for chicken model specified in *UE*, Equation 6.8, p. 166, will be used to demonstrate most of the procedures reviewed in this chapter.

Creating a residual series from a regression model:

Follow these steps to estimate the demand for chicken model (*UE*, Equation 6.8, p. 166), save the results in an equation named *EQ01*, make a residual series named *E*, and save changes to the workfile:

Step 1. Open the EViews workfile named *Chick6.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar and enter *Y C PC PB YD* in the Equation Specification: window, and click **OK**.
Step 3. Select **Name** on the equation window menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.
Step 4. To create a new series for the residuals (errors) for *EQ01*, select **Procs/Make Residual Series** on the equation window menu bar and the graphic on the right appears. Enter *E* in the Name for residual series: window, click **OK**, and a spreadsheet view of the residual series will be displayed in a new window**.**

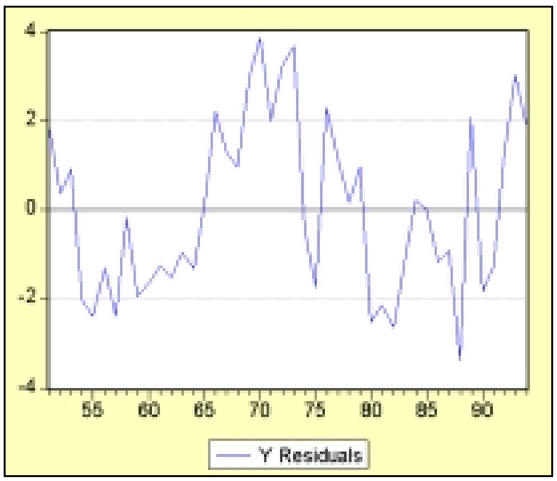


Step 5. Select **Save** on the workfile menu bar to save your changes.

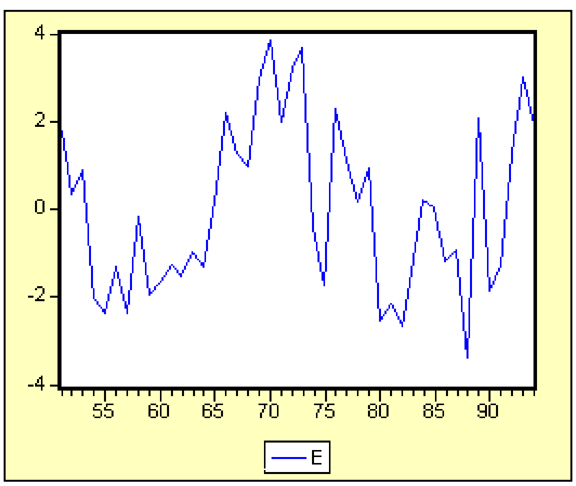
Plotting of the error term to detect serial correlation (*UE*, pp. 313-315):

Complete the section entitled Creating a residual series from a regression model before attempting this section (i.e., Equation *EQ01* and series *E* should already be present in the workfile). Follow these steps to view a residuals graph in EViews:

Step 1. Open *EQ01*, by double clicking the eq01 icon in the workfile window.
Step 2. Select **View/Actual, Fitted, Residual/Residual Graph** on the equation window menu bar to reveal the graph on the left below. Note the residual series exhibits a pattern akin to the graphs displayed in *UE*, Figure 9.1, p. 313. Thus, graphical analysis indicates positive serial correlation. **Steps 3** and **4** below show how to generate a time series plot of the same residual series *E*.



Step 3. Open the residual series named *E* in a new window by double clicking the series icon e in the workfile window to open the residual series from *EQ01* in a new window.
Step 4. Select **View/Line Graph** to reveal a time series graph of the residuals shown below.



Using regression to estimate ρ, the first order serial correlation coefficient (*UE*, Equation 9.1, pp. 311-312):[1]

Complete the section entitled Creating a residual series from a regression model before attempting this section (i.e., Equation *EQ01* and series *E* should already be present in the workfile). Follow the steps below to estimate the first order serial correlation coefficient and test for possible first order serial correlation:

Step 1. Open the EViews workfile named *Chick6.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *E C E(-1)* in the Equation Specification: window, and click **OK** to reveal the regression output shown in the graphic below. Rho (ρ) is used to symbolize the coefficient on E(-1) and it represents the first-order autocorrelation coefficient in this regression. In this case, the value of ρ is positive and significant at the 1% level (t-statistic = 3.69 and Prob value = 0.0006). It is important to note that this is not a test of serial correlation, but the value of ρ is related to the value of the Durbin-Watson d statistic discussed in the next section.[2]

Dependent Variable: E
Method: Least Squares
Date: 06/15/00 Time: 09:26
Sample(adjusted): 1952 1994
Included observations: 43 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.019618	0.258797	-0.075803	0.9399
E(-1)	0.500358	0.135451	3.694016	0.0006

R-squared	0.249713	Mean dependent var	-0.041758
Adjusted R-squared	0.231413	S.D. dependent var	1.935224
S.E. of regression	1.696593	Akaike info criterion	3.940516
Sum squared resid	118.0155	Schwarz criterion	4.022432
Log likelihood	-82.72110	F-statistic	13.64575
Durbin-Watson stat	2.105547	Prob(F-statistic)	0.000646

Step 3. Select **Name** on the equation menu bar, enter *EQ02* in the Name to identify object: window, and click **OK**.

Step 4. Select **Save** on the workfile menu bar to save your changes.

---

[1] To test for possible second order serial correlation, regress the residuals against its value lagged one period and two periods by entering *E C E(-1) E(-2)* in the Equation Specification: window, and click **OK**. To detect seasonal serial correlation in a quarterly model, regress the residuals against its value lagged four periods enter *E C E(-4)* in the Equation Specification: window, and click **OK**. Similarly, to detect seasonal serial correlation in a monthly model, regress the residuals against its value lagged twelve periods enter *E C E(-12)* in the Equation Specification: window, and click **OK**.

[2] The Durbin-Watson d statistic is approximately equal to 2(1-ρ).

Viewing the Durbin-Watson d statistic in the EViews Estimation Output window (*UE* 9.3):

Complete the section entitled Creating a residual series from a regression model before attempting this section (i.e., Equation *EQ01* should already be present in the workfile). Follow these steps to view the Durbin-Watson d test for *EQ01*:

Step 1. Open *EQ01* by double clicking the eq01 icon in the workfile window.
Step 2. Select **View/Estimation Output** on the *EQ01* menu bar to reveal the regression output shown below. The Durbin-Watson statistic is highlighted in yellow and boxed in red.[3]

Dependent Variable: Y
Method: Least Squares
Date: 06/16/00 Time: 13:06
Sample: 1951 1994
Included observations: 44

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	31.49604	1.312586	23.99541	0.0000
PC	-0.729695	0.080020	-9.118941	0.0000
PB	0.114148	0.045686	2.498536	0.0167
YD	0.233830	0.016447	14.21738	0.0000

R-squared	0.986828	Mean dependent var	43.37500
Adjusted R-squared	0.985840	S.D. dependent var	16.83854
S.E. of regression	2.003702	Akaike info criterion	4.314378
Sum squared resid	160.5929	Schwarz criterion	4.476577
Log likelihood	-90.91632	F-statistic	998.9207
Durbin-Watson stat	0.978759	Prob(F-statistic)	0.000000

Step 3. Use the Sample size printed after *Included observations:* (i.e., 44) and the number of explanatory variables listed in the *Variable* column (i.e., 3) and follow the instructions in *UE*, p. 612 to find the upper and lower critical d value in *UE*, Tables B-4, B-5 or B-6, pp. 613 - 615).

[3] The Durbin-Watson statistic is a test for first-order serial correlation. More formally, the DW statistic measures the linear association between adjacent residuals from a regression model. The Durbin-Watson is a test of the hypothesis $\rho=0$ in the specification:

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

If there is no serial correlation, the DW statistic will be around 2. The DW statistic will fall below 2 if there is positive serial correlation (in the worst case, it will be near zero). If there is negative correlation, the statistic will lie somewhere between 2 and 4. Positive serial correlation is the most commonly observed form. As a rule of thumb, with 50 or more observations and only a few independent variables, a DW statistic below about 1.5 is a strong indication of positive first order serial correlation.

Estimating generalized least squares (GLS) equations using the AR(1) method (*UE* 9.4.2):

Follow these steps to estimate the chicken demand model using the AR(1) method of GLS equation estimation.

Step 1. Open the EViews workfile named *Chick6.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C PC PB YD AR(1)* in the Equation Specification: window, and click **OK** to reveal the output below. EViews automatically adjusts your sample to account for the lagged data used in estimation, estimates the model, and reports the adjusted sample along with the remainder of the estimation output.

Dependent Variable: Y
Method: Least Squares
Date: 06/17/00 Time: 07:59
Sample(adjusted): 1952 1994
Included observations: 43 after adjusting endpoints
Convergence achieved after 14 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	26.72991	3.994877	6.691046	0.0000
PC	-0.109878	0.084948	-1.293484	0.2037
PB	0.090290	0.043806	2.061150	0.0462
YD	0.242016	0.026520	9.125644	0.0000
AR(1)	0.902894	0.063699	14.17431	0.0000

R-squared	0.995060	Mean dependent var	43.87674
Adjusted R-squared	0.994540	S.D. dependent var	16.70169
S.E. of regression	1.234162	Akaike info criterion	3.367605
Sum squared resid	57.87989	Schwarz criterion	3.572396
Log likelihood	-67.40351	F-statistic	1913.442
Durbin-Watson stat	2.159868	Prob(F-statistic)	0.000000

Inverted AR Roots	.90

The estimated coefficients, coefficient standard errors, and t-statistics may be interpreted in the usual manner. The estimated coefficient on the AR(1) variable is the serial correlation coefficient of the unconditional residuals.[4]

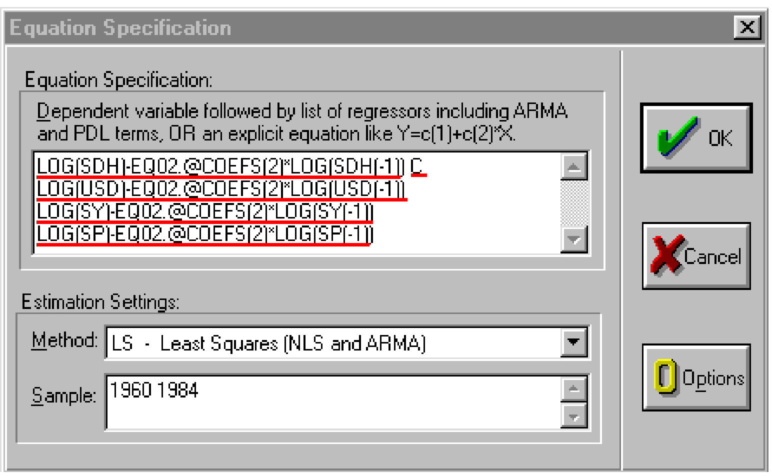
[4] Unconditional residuals are the errors that you would observe if you made a prediction of the value of using contemporaneous information, but ignoring the information contained in the lagged residual. For AR models estimated with EViews, the residual-based regression statistics—such as the, the standard error of regression, and the Durbin-Watson statistic— reported by EViews are based on the one-period-ahead forecast errors.

The most widely discussed approaches for estimating AR models are the Cochrane-Orcutt, Prais-Winsten, Hatanaka, and Hildreth-Lu procedures. These are multi-step approaches designed so that estimation can be performed using standard linear regression. EViews estimates AR models using nonlinear regression techniques. This approach has the advantage of being easy to understand, generally applicable, and easily extended to nonlinear specifications and models that contain endogenous right-hand side variables.

Estimating generalized least squares (GLS) equations using the Cochrane-Orcutt method (*UE* 9.4.2):

The Cochrane-Orcutt method is a multi-step procedure that requires re-estimation until the value for the estimated first order serial correlation coefficient converges. Follow these steps to use the Cochrane-Orcutt method to estimate the CIA's "high" estimate of Soviet defense expenditures (i.e., this is *UE*, Exercise 14, Equation 9.28, p. 342).

Step 1. Open the EViews workfile named *Defend9.wf1*.
Step 2. Follow the steps in Creating a residual series from a regression model to estimate the OLS equation *LOG(SDH) C LOG(USD) LOG(SY) LOG(SP)*, name it *EQ01*, and create a residual series for *EQ01* named *E*.
Step 3. Estimate ρ, and name it *EQ02* .
Step 4. To estimate the generalized differenced form of *UE*, Equation 9.28, select **Objects/New Object/Equation** on the workfile menu bar, enter *EQ03* in the Name to identify object: window, and enter *LOG(SDH)-EQ02.@COEFS(2)*LOG(SDH(-1)) C LOG(USD)-EQ02.@COEFS(2)*LOG(USD(-1)) LOG(SY)-EQ02.@COEFS(2)*LOG(SY(-1)) LOG(SP)-EQ02.@COEFS(2)*LOG(SP(-1))* in the Equation Specification: window. The specification should appear as in the figure below.[5] Click **OK** to view the EViews OLS output. The variable names are truncated in the EViews regression output table because they don't fit in the variable name cell. Nonetheless, the regression is correct. [6]



Step 5. To calculate the new residual series, enter the following formula in the command window: *series E = LOG(SDH)-(EQ03.@COEFS(1) + EQ03.@COEFS(2)*LOG(USD)*

---

[5] Statistical output for previously saved equations can be recalled by typing the equation name followed by a period and the reference to the specific output desired. In this case, the value for ρ from *EQ02* (recall that ρ was coefficient # 2 on the *E(-1)* term) can be recalled with the command *EQ02.@coefs(2)*. The expression *EQ02.@coefs(2)* can be used for ρ in the Equation Specification: window.
[6] The equation can be viewed by selecting **View/Representations** on the equation menu bar. The equation should read: *LOG(SDH)-EQ02.@COEFS(2)*LOG(SDH(-1)) = 0.1506791883 + 0.107961186*(LOG(USD)-EQ02.@COEFS(2)*LOG(USD(-1))) + 0.1368904004*(LOG(SY)-EQ02.@COEFS(2)*LOG(SY(-1))) - 0.000837025419*(LOG(SP)-EQ02.@COEFS(2)*LOG(SP(-1)))*. Select **View/Estimation Output** on the group window menu bar to restore the estimation output view for *EQ01*.

*+ EQ03.@COEFS(3)*LOG(SY) + EQ03.@COEFS(4)*LOG(SP))*, and press **Enter**. The phrase *"E successfully computed"* should appear in the lower left of your screen.

Step 6. Re-run *EQ02*, *EQ03* and the *series E* equation[7] in **Step 6** sequentially until the estimated ρ (i.e., the coefficient on the *E(-1)* term from *EQ02*) does not change by more that a pre-selected value such as 0.001. After 11 iterations the value for ρ converged (i.e., ρ changed from 0.957566 to 0.95758 between the 10^{th} and 11^{th} iteration).

Step 7. Convert the constant from the final version of *EQ03* by typing the following formula in the command window: *scalar BETA0=EQ03.@COEFS(1)/(1-EQ02.@COEFS(2))*, and press **Enter**. Double click beta0 the icon in the workfile window and read the value for the estimated constant in the lower left of the screen.

The final equation is *LOG(SDH) = 3.55208248072*[8] *+ 0.107961186*(LOG(USD)) + 0.1368904004*(LOG(SY)) - 0.000837025419*(LOG(SP))*. Note that this is the same equation reported in *UE*, Exercise 14, Equation 9.28, p. 342.

Exercise:

15. Follow the steps explained in the Estimating generalized least squares (GLS) equations using the Cochrane-Orcutt method section, using SDL as the dependent variable instead of SDH.

---

[7] You can re-run an equation by opening the equation in a window, selecting **Estimate** on the equation menu bar, and clicking **OK**. You can re-run the *series e* equation by clicking the cursor anywhere on the equation in the command window and hitting **Enter** on the keyboard.

[8] The number *3.55208248072* was computed in **Step 7** above.