Chapter 2: Ordinary Least Squares

In this chapter:

Ordinary Least Squares (OLS) regression is the core of econometric analysis. While it is important to calculate estimated regression coefficients without the aid of a regression program one time in order to better understand how OLS works (see *UE*, Table 2.1, p.41), easy access to regression programs makes it unnecessary for everyday analysis.[1] In this chapter, we will estimate simple and multivariate regression models in order to pinpoint where the regression statistics discussed throughout the text are found in the EViews program output. Begin by opening the EViews program and opening the workfile named *htwt1.wf1* (this is the file of student height and weight that was created and saved in Chapter 1).

Running a simple regression for weight/height example (*UE* 2.1.4):

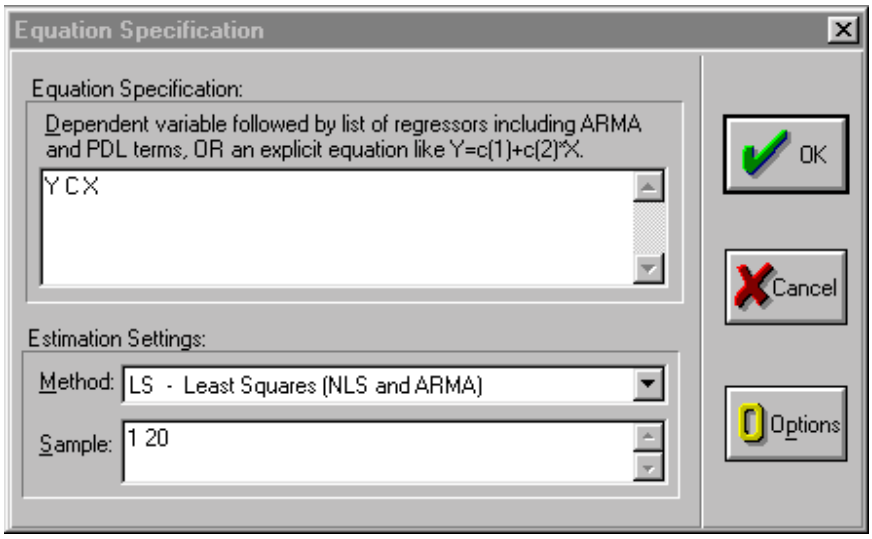
Regression estimation in EViews is performed using the equation object. To create an equation object in EViews, follow these steps:

Step 1. Open the EViews workfile named *htwt1.wf1* by selecting **File/Open/Workfile** on the main menu bar and click on the file name.

Step 2. Select **Objects/New Object/Equation** from the workfile menu.[2]

Step 3. Enter the name of the equation (e.g., *EQ01*) in the Name for Object: window and click **OK**.

Step 4. Enter the dependent variable weight (*Y*), the constant (*C*) and the independent variable height (*X*) in the Equation Specification: window (see figure on right). It is important to enter the dependent variable first (*Y* in this case).



[1] *Using Econometrics, A Practical Guide* (fourth edition), by A. H. Studenmund will be referred to as (*UE*) when referenced in this guide.

[2] Alternately, select **Quick/Estimate Equation** from the main menu. If this method is used, the equation must be named to save it. Click **Name** on the equation menu bar and enter the desired name and click **OK**.

Step 5. Select the estimation **Method** {LS - Least Squares (NLS and ARMA)}. This is the default that will be used most of the time.

Step 6. The workfile sample range is automatically entered but it can be changed if another sample range is desired. Click **OK** to view the EViews Least Squares regression output table.

Step 7. To save changes to your workfile, click **Save** on the workfile menu bar.

Contents of the EViews equation window:

Most of the important statistical information relating to a regression is reported in the EViews equation window (see the figure below). General information concerning the regression is printed in the top few lines, the coefficient statistics are reported in table format (middle five columns), and summary statistics are printed in table format (bottom four columns) of the equation window.

General Information Printed in the Top Portion of the Equation Output: The first five or six lines identify (see arrow in figure on right):

Line 1. Name of the dependent variable.

Line 2. Regression method used.

Line 3. Date & time the regression was executed.

Line 4. Sample range used in the regression.

Line 5. Number of observations included in your regression.

Line 6. Number of excluded observations.

Line 6 is not reported in this case because no observations are excluded (i.e., no variable, included in the regression, has missing observations or is lagged with data not available for the pre-sample period).

Dependent Variable: Y
Method: Least Squares
Date: 06/24/00 Time: 09:02
Sample: 1 20
Included observations: 20

General information concerning the regression

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	103.3971	9.342100	11.06786	0.0000
X	6.377093	0.883732	7.216091	0.0000

R-squared	0.743121	Mean dependent var	169.4000
Adjusted R-squared	0.728850	S.D. dependent var	16.32692
S.E. of regression	8.501763	Akaike info criterion	7.213064
Sum squared resid	1301.039	Schwarz criterion	7.312637
Log likelihood	-70.13064	F-statistic	52.07197
Durbin-Watson stat	1.451314	Prob(F-statistic)	0.000001

Coefficient Results: Key information regarding the estimated regression coefficients is reported in a table displayed in the middle of the regression output (see area highlighted in yellow). The first column identifies each variable (*C* for the constant and *X* for height) and the second column reports the estimated coefficient values (i.e., $\hat{\beta}_o$ & $\hat{\beta}_1$). Note that the estimated coefficients are the same as those printed in (*UE*, Equation 1.21, pp. 20 & 21). The data printed in columns (3) - (6) are the standard error of the coefficient, the t-statistic and the probability respectively. We will not discuss these now, but they will be very important when *Hypothesis Testing* is discussed in Chapter 5.

Summary Statistics: Key summary statistics are reported in four columns below the equation output (see area outlined in **red**). Each will be defined and the page reference to where it is discussed in *UE* is identified in parenthesis.

1. R^2: coefficient of determination is the fraction of the variance of the dependent variable explained by the independent variables (p. 50);
2. $\bar{R}^2$: adjusted R^2 (p. 51);
3. Standard Error of the Regression (S.E. of regression): called Standard Error of Estimate (SEE) in *UE* (p. 105t & 392t);
4. Sum of squared resid: OLS selects the value of the coefficients to minimize this (pp. 37–38);
5. Log likelihood: useful in hypothesis testing;
6. Durbin-Watson stat: test statistic for serial correlation in the residuals (p.324);
7. Mean dependent var: measure of central tendency for the dependent variable (pp. 522-526);
8. S.D. dependent var: measure of dispersion (Standard Deviation) for the dependent variable (pp. 526-529);
9. Akaike info criterion: used in model selection (pp. 195-197);
10. Schwarz criterion: used in model selection (pp. 195-197);
11. F-statistic: tests the hypothesis that all of the slope coefficients (excluding the constant, or intercept) in a regression are zero (pp. 142–145);
12. Prob(F-statistic): (pp. 144–145).

Multivariate Regression (2.2.3):

Multivariate regression is executed the same as simple regression in EViews and the output is identical. In this section we create a workfile, import data from a spreadsheet and estimate a multivariate regression model using the Beef example (*UE*, p. 45).

Creating a workfile for the demand for beef example (*UE*, Table 2.2, p. 45):

Note that Table 2.2 reports annual data for the price and quantity of beef and disposable personal income for the period 1960 through 1987. To create an EViews workfile:

Step 1. Select **File/New/Workfile** on the main menu.
Step 2. Set the Workfile frequency: to *Annual*.
Step 3. Enter the Start date: (*1960*) and End date: (*1987*).
Step 4. Click **OK**.

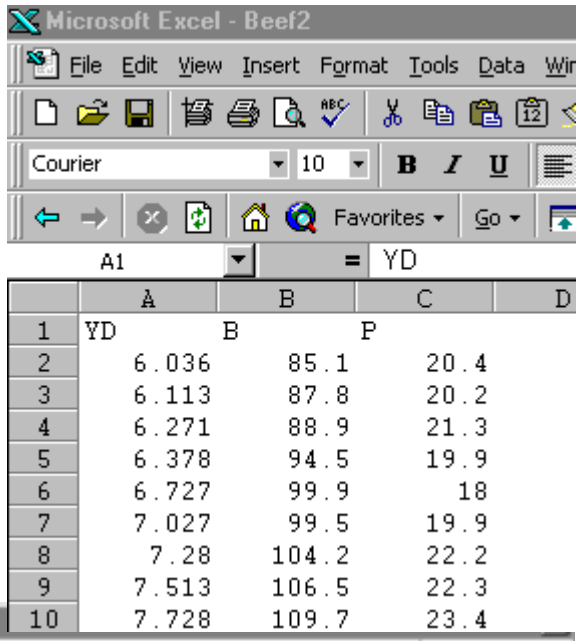
Importing data from a spreadsheet file named *Beef2.xls*:

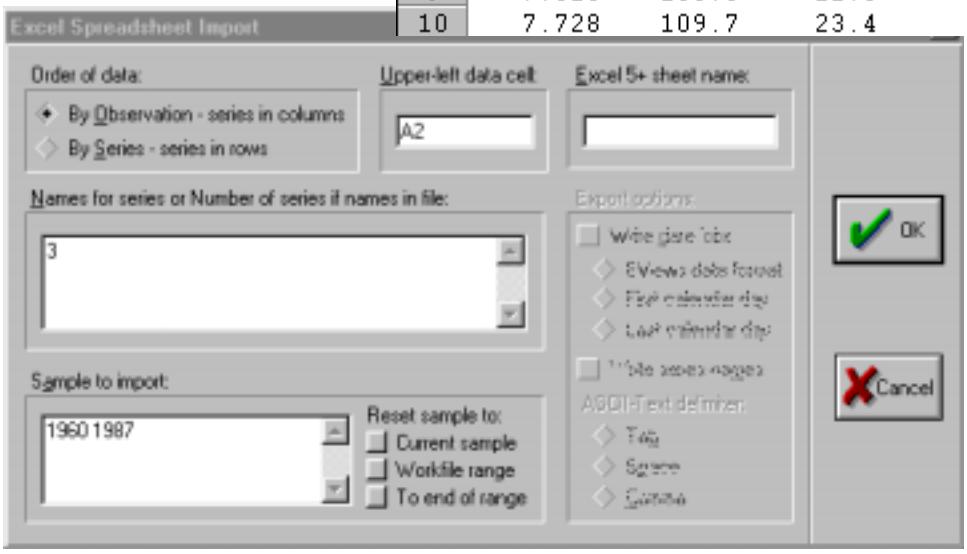
Once the workfile has been created, it is a simple matter to import data from another file. However, you must first know the location of the data in the spreadsheet file.

To view the location of the data in the spreadsheet file:

Step 1. Open the Excel program and then open the Excel spreadsheet file named *Beef2.xls*, located on the Addison Wesley Longman web site. You can get there by clicking this *http://occ.awlonline.com/bookbind/pubbooks/studenmund_awl/* URL. Click on the *Student Resources/Data Sets/Data Sets-Excel*. You must be connected to the internet for this link to work.
Step 2. Note that the first data series starts in cell A2 and that the data are in three contiguous columns (see the figure below).

Follow these steps to import the data from the *Beef2.xls* Excel file into your new EViews file:





Step 1. Close the spreadsheet file (two programs cannot access the same file at the same time).

Step 2. Click **Procs/Import/Read Text-Lotus-Excel** on the workfile menu bar[3].

Step 3. Select the drive and folder location in the Look in: window.

Step 4. Select *Excel.xls* in the Files of type: window.

Step 5. Double click on the file named *Beef2.xls*.

Step 6. Fill in *A2* for the upper left data cell and *3* for the number of series (note that when you enter the number of series, EViews will enter the names of the series that are printed in the row above each data series).

Step 7. The sample range is set to the workfile sample by default. The screen should look like the figure on the right.

Step 8. Click **OK** to complete the import process. If you get an error message that reads 'Unable to open file …' it means that you probably have not closed the file in Excel. If the error message does not display, the data was successfully imported.

Using EViews to estimate a multiple regression model of beef demand (*UE* 2.2.3):

To Regress Beef Demand (*B*) on the Constant (*C*), the Price of Beef (*P*) and Per Capita Disposable Income (*Yd*):

Step 1. Open the EViews workfile named *Beef2.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar and enter *B C P Yd* in the Equation Specification: window. Do not change the default settings for Method and Sample.

[3] If you know the location and name for the spreadsheet file, you can skip steps 3–5. Instead, fill in the **File name** box and click **OK**.

Step 3. Click **OK** to get the regression results shown in the table on the right. Check the coefficients in column 2 of the EViews, Least Squares output table (highlighted in yellow) with the results reported in *UE*, Equation 2.10, p. 44.

Dependent Variable: B
Method: Least Squares
Date: 07/07/00 Time: 09:02
Sample: 1960 1987
Included observations: 28

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	37.53605	10.04020	3.738575	0.0010
P	-0.882623	0.164730	-5.357981	0.0000
YD	11.89115	1.762162	6.748045	0.0000

R-squared	0.658030	Mean dependent var	106.6500
Adjusted R-squared	0.630672	S.D. dependent var	10.00561
S.E. of regression	6.080646	Akaike info criterion	6.549056
Sum squared resid	924.3564	Schwarz criterion	6.691792
Log likelihood	-88.68678	F-statistic	24.05287
Durbin-Watson stat	0.292597	Prob(F-statistic)	0.000001

Describing the Overall Fit of the Estimated Model (*UE* 2.4):

R^2, the coefficient of determination and $\overline{R}^2$, the adjusted R^2, are located right under the coefficients column of the table printed in the middle of the EViews regression output (outlined in red). In this case, $R^2 = 0.66$ and $\overline{R}^2 = 0.63$.

Exercises:

4. Follow the steps described in Chapter 1 of the EViews guide to create an EViews workfile and enter the data into an EViews workfile. Follow the steps described in Running a simple regression for weight/height example of the EViews guide to regress *per capita income* as a function of *percent of labor in agriculture* on farms.