A Guide to Using EViews with
Using Econometrics: A Practical Guide

Written By

R. R. Johnson[*]
University of San Diego

Introduction

In the Introduction:
1. Purpose of this guide
2. EViews help
3. EViews basics and objects
4. Mathematical expressions in EViews
5. Areas in EViews main window

Purpose of this guide:

Econometrics is defined by A. H. Studenmund[1] to be "the quantitative measurement and analysis of actual economic and business phenomena." This definition highlights the importance of learning econometrics by working through examples. This guide demonstrates how to use EViews to complete the econometric analysis illustrated in the text. EViews output can be copied and pasted into word processing files to facilitate the research report writing process.

Each chapter in this guide corresponds to a chapter in *Using Econometrics: A PracticalGuide (UE)*, by A. H. Studenmund. All of the econometric processes examined in a chapter are reproduced in the corresponding chapter of this guide, using EViews whenever a data set is provided.[2] This guide is a do-it-yourself manual and students should be able to reproduce the econometric analysis described in *UE*, without further assistance from the instructor. Best results are achieved when the text chapter is read first and then the operations are performed as you read the guide. Most procedures described in this guide are explained in a step-by-step manner so that even the novice user should be able to follow. The reader is advised to consult the owner's manual or the help function in EViews in cases where: a deeper understanding of EViews' functions is desired, a description of how EViews performs the calculations is desired, or references to the source of the econometric theory applied by EViews is sought.

---

[*] I would like to thank A. H. Studenmund for his helpful comments and a special thanks to Pat Johnson for her superb editorial assistance and for making sure that the instructions in this guide are clear and easy to follow. Your feedback is very valuable - send your comments about this site and this guide to mailto:Johnson@acusd.edu.

[1] A. H. Studenmund, *Using Econometrics, A PracticalGuide* (fourth edition), Addison Wesley, 2000, p. 3.

[2] For a variety of reasons, data sets for some of the examples presented in the text are not available.

EViews help:

For additional information, open the EViews program and select **Help/EViews Help Topics…** and a list of help categories is revealed (see the graphic below). Each help category contains a list of subtopics that are revealed by double clicking the book icon in front of the category. Other topics and subtopics can be viewed by scrolling down and clicking the subject of interest. If you know the specific subject matter, you can view the **Contents** listed in alphabetical order or use **Find** to locate the topic of interest. EViews uses standard Windows **Help**, the on-line manual is fully searchable and hypertext linked. The **Help** system contains updates to the documentation that were made after the manuals went to press. In addition to the information provided in the manuals and the help system, go to http://www.eviews.com on the World Wide Web and review the **Frequently Asked Questions (FAQ)** section, where you can find answers to common questions about installing, using, and getting the most out of EViews.
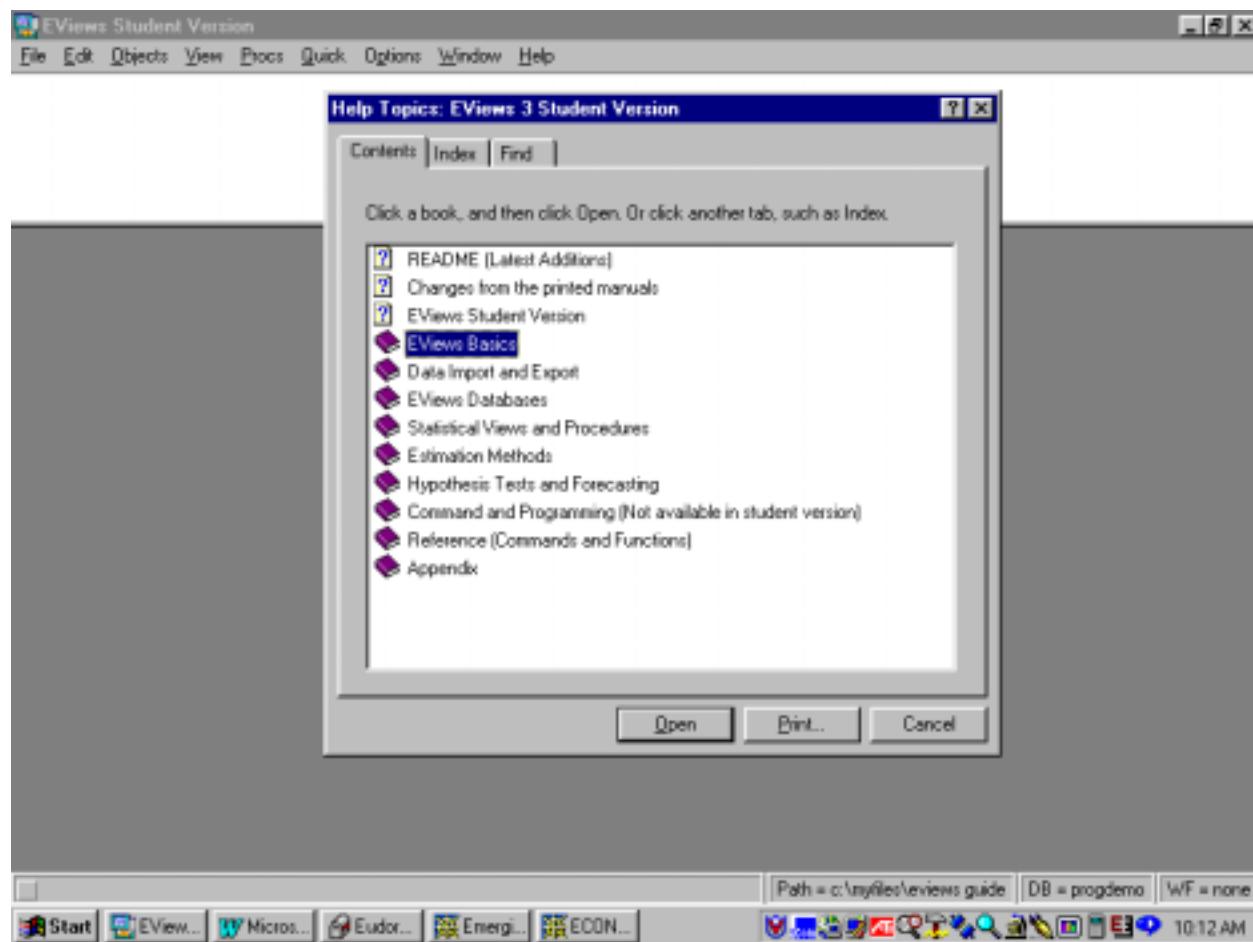
EViews basics and objects:

The EViews program is designed around the concept of objects. Objects are collections of related information and operations that are bundled together in an easy-to-use unit. Virtually all of your work in EViews will involve using and manipulating various objects. Think of an object as a filing cabinet or organizer for the item with which you are working. The most important object in EViews is the workfile and your first step in any project will be to create a new workfile or to load an existing workfile into memory. Each object consists of a collection of information related to a particular area of analysis. Objects associated with a particular concept are said to be of a given type, where the type name is used to identify the subject of analysis. Associated with each type of object is a set of views and procedures which can be used in association with the information contained in the object. For example, a series object is a collection of information related to a set of observations on a particular variable. An equation object is a collection of information related to the relationship between a collection of variables. Since an equation object contains all of the information relevant to an estimated relationship, you can move freely between a variety of equation specifications simply by choosing to work with a different equation object. You can examine results, perform hypothesis and specification tests, or generate forecasts at any time. Managing your work is simplified since only a single object is used to work with an entire collection of data and results. **Tip:** Name all objects that you think you want to view later. Then the object can be recalled by double clicking the named object in the workfile.

The most fundamental objects in EViews are workfiles, series, and equation objects. There are, however, a number of other objects that serve special functions. A list of EViews objects includes: Coefficient Vector, Databases, Equation, Graph, Group, Model, Pool (Time Series / Cross-Section), Sample, Series, State Space, System, SYM (Symmetric Matrix), Table, Text, VAR (Vector Autoregression), Vector/Row, and Vector Scalar. All objects, except workfiles and databases, have their own icons that are displayed in the workfile window (see object types below along with their respective icons). When a new workfile is created, two objects are displayed in the workfile window – the Coefficient Vector (filled with zeros) and the residual series (filled with NA's).

Object Types

The most fundamental objects in EViews are series and equation objects. There are, however, a number of different types of objects, each of which serves a unique function. Most objects are represented by a unique icon which is displayed in the object container window:

Coefficient Vector	Series
Equation	State Space
Graph	System
Group	SYM (Symmetric Matrix)
Matrix	Table
Model	Text
Pool (Time Series / Cross-Section)	VAR (Vector Autoregression)
Sample	Vector/Row Vector
Scalar	

Despite the fact that they are also objects, object containers do not have icons since they cannot be placed in other object containers—thus, workfiles and databases do not have icons since they cannot be placed in other workfiles or databases.

To create an object: select **Objects/New Object** from the main menu or workfile menu, click on the type of object that you want to create, provide a name and then click on **OK**. For some object types, another dialog box will open, prompting you to describe your object in more detail. For most objects, however, the object window will open immediately. A workfile must be open before an object can be created. The process of creating a workfile and objects will be explained later in this guide.

Mathematical expressions in EViews:

EViews contains an extensive library of built-in operators and functions that allow you to perform complicated mathematical operations with your data using just a few keystrokes. In addition to supporting standard mathematical and statistical operations, EViews provides a number of specialized functions for automatically handling the leads, lags, and differences that are commonly found in time series data. All of the operators described below may be used in expressions involving series and scalar values. When applied to a series expression, the operation is performed for each observation in the current sample.

EViews follows the usual order in evaluating expressions from left to right, with operator precedence order (from highest precedence to lowest):

1. ^
2. *, /
3. +, subtraction (-)
4. <, >, <=, >=, =
5. and, or

For a list and description of all of the operators and special functions available in EViews, click on: **Help/Function Referenc*e***.

The table below explains the function of each mathematical operator used by EViews.

Expression	Operator	Description
+	add	x+y adds the contents of X and Y
–	subtract	x–y subtracts the contents of Y from X
*	multiply	x*y multiplies the contents of X by Y
/	divide	x/y divides the contents of X by Y
^	raise to the power	x^y raises X to the power of Y
>	greater than	x>y takes the value 1 if X exceeds Y, and 0 otherwise
<	less than	x<y takes the value 1 if Y exceeds X, and 0 otherwise
=	equal to	x=y takes the value 1 if X and Y are equal, and 0 otherwise
<>	not equal to	x<>y takes the value 1 if X and Y are not equal, and 0 if they are equal
<=	less than or equal to	x<=y takes the value 1 if X does not exceed Y, and 0 otherwise
>=	greater than or equal to	x>=y takes the value 1 if Y does not exceed X, and 0 otherwise
and	logical and	x and y takes the value 1 if both X and Y are nonzero, and 0 otherwise
or	logical or	x or y takes the value 1 if either X or Y is nonzero, and 0 otherwise

Areas in EViews main window:
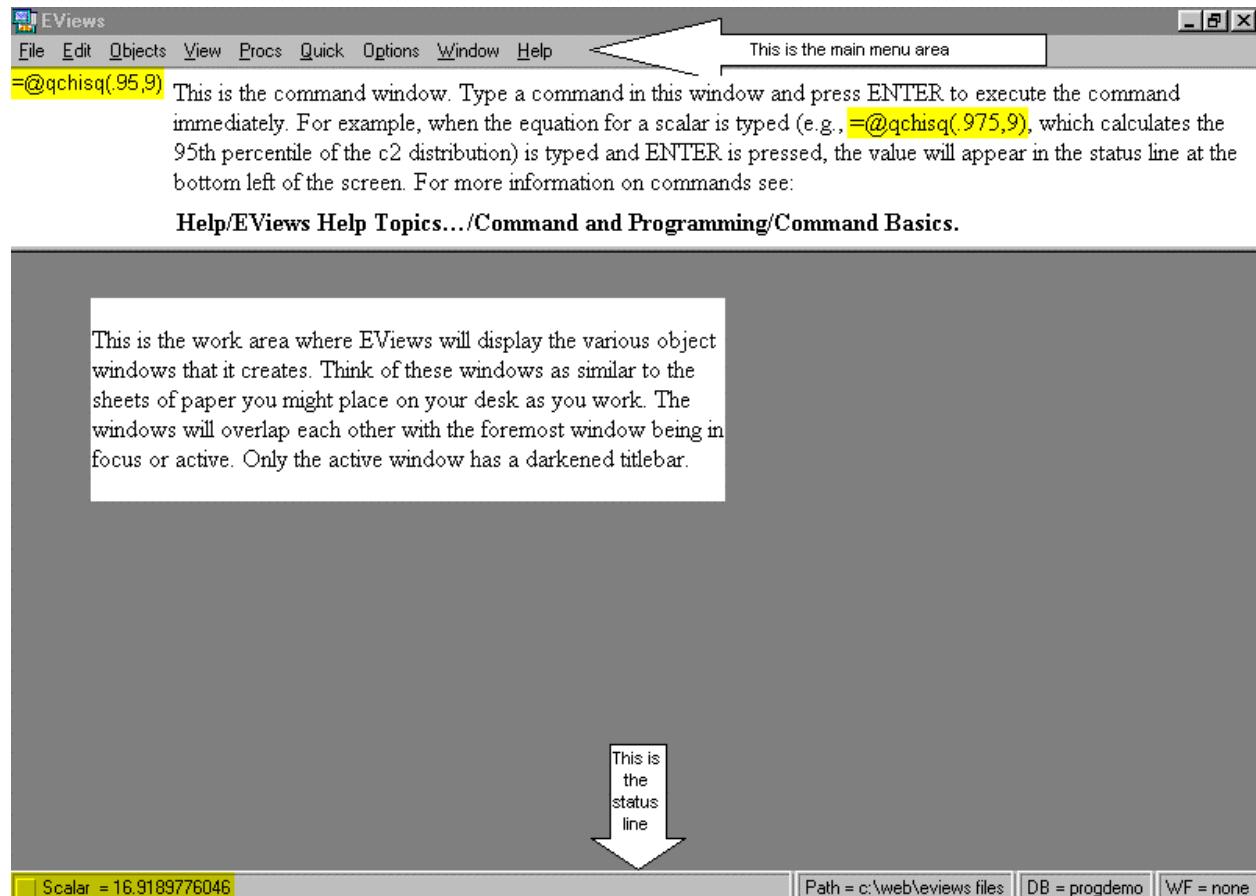
Refer to the graphic above as you read a description of the various areas of the EViews window below. References to these areas will be made throughout this guide.

The title bar: The title bar, labeled EViews Student Version, is at the very top of the main window.

The main menu: Just below the title bar is the main menu. If you move the cursor to an entry in the main menu and click on the left mouse button, a drop-down menu will appear. Clicking on an entry in the drop-down menu selects the highlighted item. Some of the items in the drop-down menu may be listed in black and others in gray. In menus, black items may be executed while the gray items are not available.

The command window: Below the main menu bar is an area called the command window where EViews commands are typed. The command is executed as soon as you hit **ENTER**.

The status line: At the very bottom of the window is a status line, which is divided into four sections. The left section will sometimes contain status messages sent to you by EViews. These status messages can be cleared manually by clicking on the box at the far left of the status line. The next section shows the default directory that EViews will use to look for data and programs. The last two sections display the names of the default database and workfile.

The work area: The area in the middle of the window is the work area where EViews will display the various object windows that it creates. Think of these windows as similar to the sheets of paper you might place on your desk as you work. The windows will overlap each other with the foremost window being in focus or active. Only the active window has a darkened titlebar.

Chapter 1: An Overview of Regression Analysis

In this chapter:

1. A simple example of regression analysis (*UE* 1.4):
   a) [Creating an EViews workfile](#)
   b) [Entering data into an EViews workfile](#)
   c) [Creating a group in EViews](#)
   d) [Graphing with EViews](#)
   e) [Generating new variables in EViews](#)
2. [Exercises](#)

Section 1.4 describes how a weight guesser can use regression analysis to make better guesses about a person's weight (*Y*) based only on a person's height (*X*). The data for this example are printed in *UE*[1], Table 1.1, p. 20, and available in the EViews workfile named *htwt1.wf1*. However, since some data sets may not be in EViews file format, the process of creating an EViews workfile and entering data into the new workfile will reviewed here. The process of creating an EViews workfile and importing data into the new workfile from an Excel file will be reviewed in Chapter 2.

Creating an EViews workfile:

To create a new workfile for the weight (*Y*) versus height (*X*) data set printed in columns (2) and (3) of (*UE*, Table 1.1, p. 20):

Step 1. Select **File/New/Workfile** on the EViews main menu bar.

Step 2. Set the Workfile frequency: to Undated or irregular.

Step 3. Since there are 20 observations in the data set, set the Start observation to 1 and End observation to 20 (refer to areas highlighted in yellow in the figure on the right).[2]

Step 4. Once you have selected the appropriate frequency and entered the information for the workfile range, click **OK**. EViews will create an untitled workfile, and will display the workfile window in the main work area of the EViews screen. The workfile window displays two pairs of numbers highlighted in yellow: one for the Range: of data contained in the workfile, and the second for the current workfile Sample:. Both the workfile range and sample can be
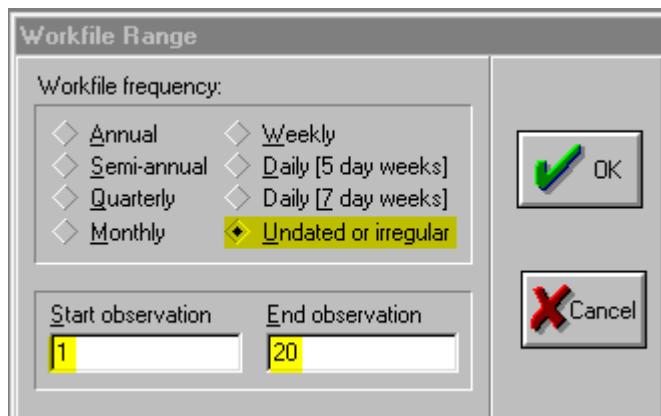
---

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

[2] For monthly, quarterly, or annual time series data, with observations ranging from the beginning of 1980 to the end of 1996, select from the following options: for monthly data, set the Workfile frequency: to *monthly*, and specify the Start date, 1980:01, and the End date, 1996:12. For quarterly data, set the Workfile frequency: to *quarterly*, and specify the Start date, 1980:1, and the End date, 1996:4. For annual data, set the Workfile frequency: to *annual*, and specify the Start date, 1980, and the End date, 1996. See **Help/Contents/EViews Basics/Workfile Basics** for more information about creating and using workfiles in EViews.

changed after the workfile has been created. Note that all new workfiles will contain two objects boxed in red: the coefficient vector named *c* and the residual series named *resid*. (see the figure below).

Step 5. To save your workfile, select **Save** on the workfile menu bar or **File/Save** or **File/Save As** on the main menu bar and enter *htwt1.wf1* in the File <u>n</u>ame: window, select the destination drive in the Save <u>in</u>: window and click **OK**. It is advisable to save the workfile after it is created and frequently thereafter.
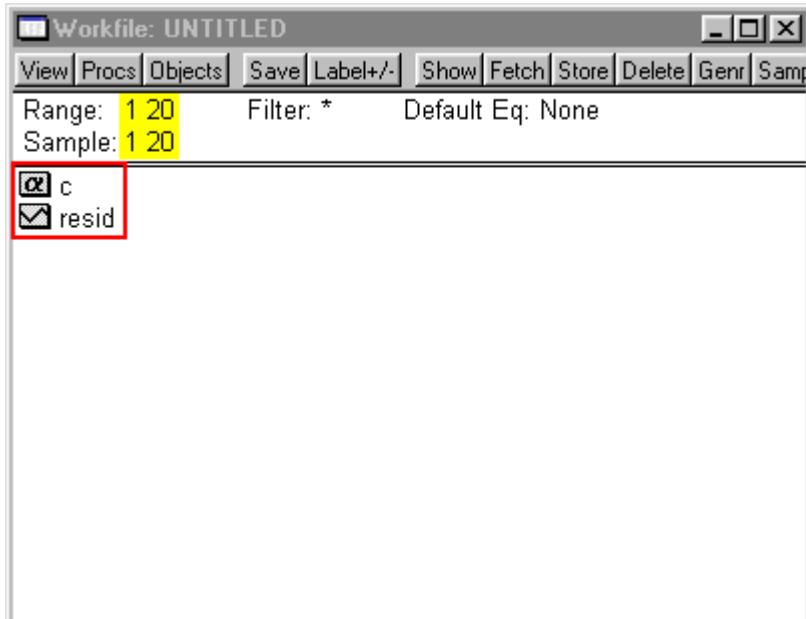
Entering data into an EViews workfile:

Complete the section entitled <u>Creating an EViews workfile</u> before attempting this procedure. To enter the data for height (inches above 5') and weight from columns (2) and (3) of Table 1.1, (*UE*, p. 20) into the newly created workfile, follow the steps below.

Step 1. To create a new series for the weight (*Y*) variable, select **Objects/New Object/Series** from the main menu or the workfile menu, enter *Y* in the Name for Object: window and click **OK**. All of the observations in the series will be assigned the missing value code 'NA'.[3]

Step 2. To enter data into the newly created series, double click the series in the workfile window and click **edit+/-** on the series window menu bar. The numbers from the table can be entered to replace the NA's in the spreadsheet, pressing **Enter** after each entry. After the numbers have been entered, click **edit+/-** on the series window menu bar to save the changes and exit the edit function. The series window can be closed by clicking the ☒ button in the upper right corner of the series window.

Step 3. Repeat the process for the height (*X*) variable.

Step 4. To save changes to your workfile, click **Save** on the workfile menu bar.

Most data are available in spreadsheet file or ASCII text file format, which can be imported directly into the workfile. The procedure for importing the weight and height data from an ASCII text file will be explained in <u>Chapter 2</u>. The edit procedure described above is mostly used to adjust data series after they are imported. In cases where expressions can be used to assign values for the series, click on **Quick/Generate Series** or click on **Genr** on the workfile window menu bar and enter the expression defining the series. For more information on how to generate a series using expressions, see **Help/Contents/EViews Basics/Working with Series**.

---

[3] A series can also be created by clicking on **Quick/Generate Series** on the main menu bar or by clicking on **Genr** on the workfile window menu bar and enter the expression *Y = NA* or any other mathematical expression.

Creating a group in EViews:

EViews provides specialized tools for working with groups of series. Follow these steps to create a group object containing the *Y* & *X* series:

Step 1. Open the EViews workfile named *htwt1.wf1* by selecting **File/Open/Workfile** on the main menu and click on the file name.
Step 2. To create a group object for the *X* and *Y* variables, hold down the **Ctrl** button and click on *X* and *Y* and then select **Show** from the workfile toolbar.
Step 3. To name a group, click **Name** or **Object/Name** on the group menu bar and enter *GROUP01* in the Name to identify object: window.[4] Some features of groups are described in the footnote at the bottom of this page.[5]
Step 4. To save changes to your workfile, click **Save** on the workfile menu bar.

Graphing with EViews:

The weight guesser in section 1.4 (*UE*) hypothesized a positive relationship between weight and height. As a first step, plot *Y* against *X* in order to visually confirm the hypothesis.

Step 1. Open the EViews workfile named *htwt1.wf1*.
Step 2. To plot *Y* against *X*, open the two series in a group window (enter *X* before *Y* because EViews specifies the first variable selected in a group as the "X" axis variable and the second as the "Y" axis variable).
Step 3. Select **View/Graph/Scatter/Scatter with Regression** and click **OK** to reveal Figure 1a. EViews shows a graph with *Y* on the vertical axis and *X* on the horizontal axis as well as a line representing the regression equation between the two variables (i.e., plots *UE*, Equation 1.21, p. 21). As you can see in Figure 1a, a positive relationship between *Y* and *X* is visually confirmed but the relationship is distorted because EViews uses **Optimal-Linear Scaling** as the graph default setting.

Figure 1a

Step 4. To change the graph scaling, click the right mouse button anywhere in the graph and select **Options**.
Step 5. Change the **Graph Scaling** option to **Linear - Force Through Zero**.

---
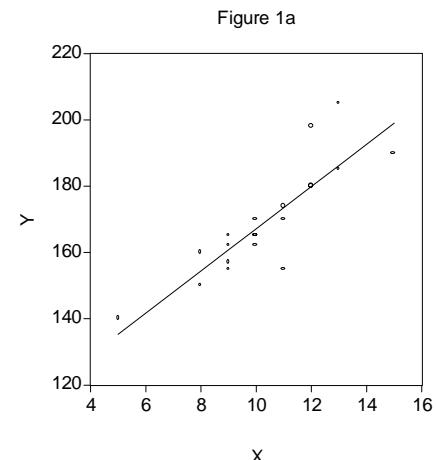
[4] You must name a group object if you want to keep its results. Unnamed objects are labeled "UNTITLED" and the results are lost when the object window or the workfile is closed. To name a group, click **Name** on the group menu bar and enter the name in the Name to identify object: window. Once named, a 'group object' is saved with the workfile and can be viewed by double clicking its icon (i.e., 🅖 group01) in the workfile window.

[5] **Features of Groups:** A group is simply a list of series identifiers, not a copy of the data in the series. If you change the data for one of the series in the group, you will see the changes reflected in the group. If you delete a series from the workfile, it will disappear from any group that included the series. If the deleted series is the only series in a group, the group will also be deleted. Renaming a series changes the reference in every group containing the series. Groups, like other EViews objects, contain their own views and procedures. A detailed description can be found in **Help/Contents/Statistical Views and Procedures/Group Views and Procs**.

Step 6. Click **OK** to get the graph in Figure 1b.

Note that the graph no longer gives a distorted view of the relationship between *Y* and *X*. The equation represented in Figures 1a & 1b has the following formula (see *UE* Equation 1.21, p. 21): *Estimated weight (Y) = 103.40 + 6.38 * Height in inches above five feet (X)*

Generating new variables in EViews:

Follow the steps below to generate the data in columns (4) through (7) of *UE*, Table 1.1, p. 20:

Figure 1b
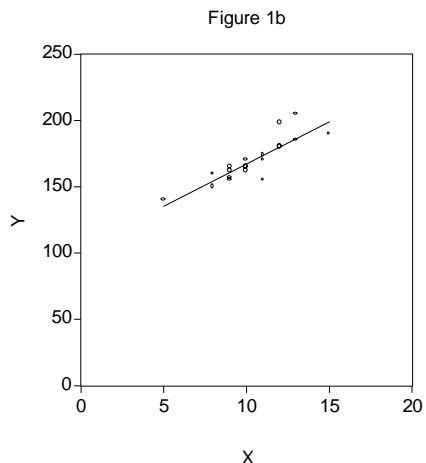
Step 1. Open the EViews workfile named *htwt1.wf1*.

Step 2. To generate a new variable named *predicted*, with the data printed in column (4) of *UE*, Table 1.1, click **Genr** on the workfile menu and enter the formula (*predicted=103.4+6.38*X*) in the Enter Equation: window. Click **OK** and a new variable named *predicted* appears in the workfile window with the values in column (4) of Table 1.1.

Step 3. To generate a new variable named *residual*, with the data printed in column (5) of Table 1.1, click **Genr** on the workfile menu and enter the formula (*residual=Y-predicted*) in the Enter Equation: window. Click **OK** and a new variable named *residual* will appear in the workfile window with the values in column (5) of Table 1.1.

Step 4. To generate a new variable named *gainorloss*, with the data printed in column (6) of Table 1.1, click **Genr** on the workfile menu and enter the formula *gainorloss=.5*(residual>=-10 and residual<=10)-.6*(residual<-10 or residual>10)* in the Enter Equation: window.[6] Click **OK** and a new variable named *gainorloss* will appear in the workfile window with the values in column (6) of Table 1.1.

Step 5. Create a group with *x, y, predicted, residual,* and *gainorloss* in order. Click **View/Spreadsheet** to view the data in Table 1.1.

Step 6. Type *=@sum(gainorloss)* in the command window, press **Enter**, view the net amount gained on the status line in the lower left of the screen (i.e., Scalar = 6.7).

---

[6] The conditional value typed within a set of parenthesis is equal to one when the condition is true and zero if the condition is false. Thus, the condition *(residual>=-10 and residual<=10)* is equal to one if the residual falls within the range -10 to +10 and zero if it is outside of this range. The condition *(residual<-10 or residual>10)* is equal to one if the residual is less than -10 or greater than +10 and zero if it falls within the range -10 to +10.

13.

    a.

    b.

    c.

    d.   The data for this exercise are available in an EViews file named *help1.wf1*. However, it is a good idea to take this opportunity to make sure that you understand how to create a workfile and enter data. If you want to skip these steps, open the EViews workfile named *help1.wf1* and skip steps 1 through 6 below.

Follow these steps to create a new workfile for the help-wanted advertising index (*hwi*) and unemployment rate (*ur*) data set printed in the table accompanying exercise 13.

Step 1. Select **File/New/Workfile** on the EViews main menu bar. Set the workfile frequency to *Quarterly*. Enter *1962:1* for the Start date: and *1967:4* for the End date: (see footnote 2). Click **OK** to open a new untitled workfile.

Step 2. To save your workfile, select **Save** on the workfile menu bar or **File/Save** or **File/Save As** on the main menu bar and enter *help1.wf1* in the File name: window, select the destination drive in the Save in: window, and click **OK**.

Step 3. To create a new series for the help-wanted advertising index (*hwi*) variable, select **Objects/New Object/Series** from the main menu or the workfile menu, enter *hwi* in the Name for Object: window and click **OK**. All of the observations in the series will be assigned the missing value code 'NA'.

Step 4. To enter data into the newly created series, double click the series in the workfile window and click **edit+/-** on the series window menu bar. The numbers from the table can be entered to replace the NA's in the spreadsheet, pressing **Enter** after each entry. After the numbers have been entered, click **edit+/-** on the series window menu bar to save the changes and exit the edit function. The series window can be closed by clicking the ☒ button in the upper right corner of the series window.

Step 5. Repeat the process for the unemployment rate (*ur*) variable.

Step 6. To save changes to your workfile, click **Save** on the workfile menu bar.

The process of running a regression is explained, in detail, in Chapter 2, but a short description of the process is presented here. Follow these steps to run the regression between *hwi* and *ur*:

Step 1. Open the EViews workfile named *Help1.wf1*.

Step 2. Select **Objects/New Object/Equation** and click **OK** (or **Quick/Estimate Equation** from the main menu).

Step 3. Enter the dependent variable (*hwi*), the constant (*c*) and the independent variable (*ur*) in the Equation Specification: window that appears, using spaces between each term. It is important to enter the dependent variable first (*hwi* in this case).

Step 4. Select the estimation method {LS - Least Squares (NLS and ARMA)}. EViews uses this as the default setting because it is selected most of the time.

Step 5. The workfile sample range is automatically entered but it can be changed if another sample range is desired.

Step 6. Click **OK** when finished to reveal the regression output generated by EViews.

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).
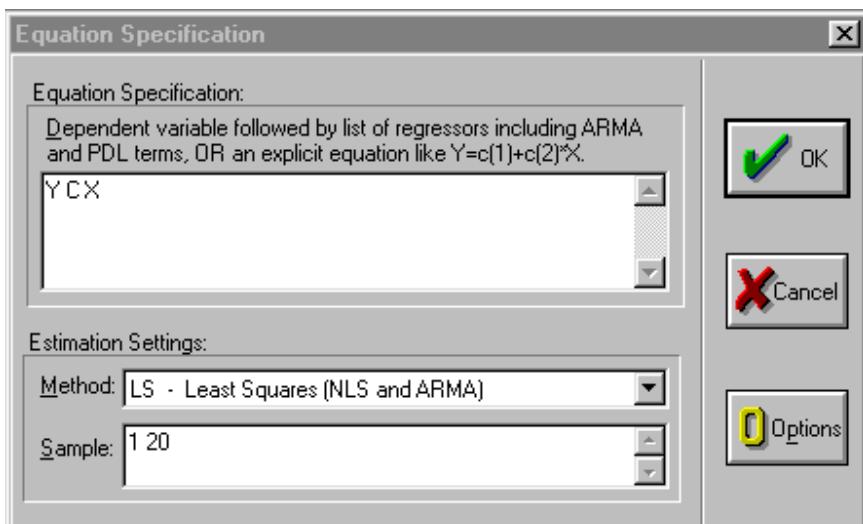
---

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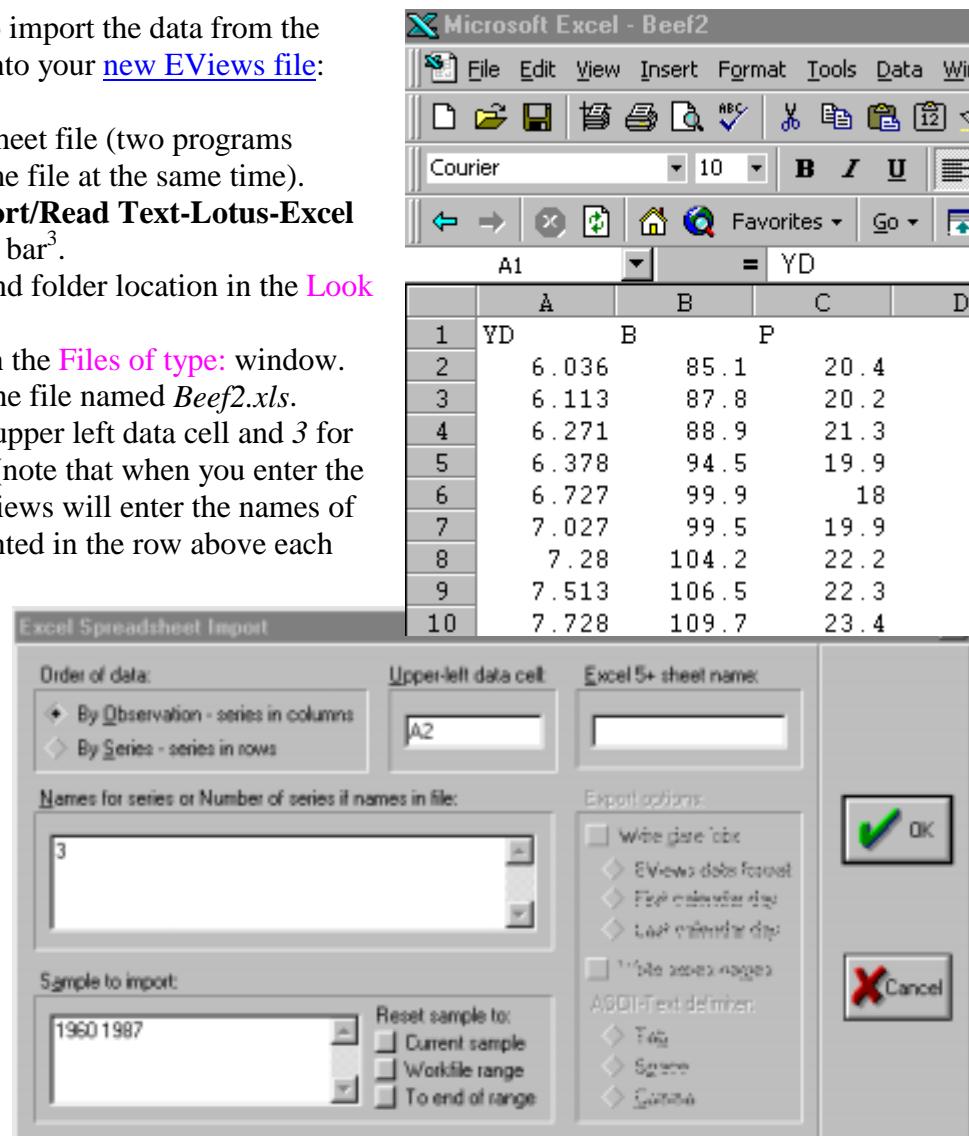
Using EViews to estimate a multiple regression model of beef demand (*UE* 2.2.3):

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 <mark>yellow</mark>) with the results reported in *UE*, Equation 2.10, p. 44.

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

```
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

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.

Chapter 3: Learning To Use Regression Analysis

In this chapter:
1. Displaying the spreadsheet view of a group of variables (*UE*, Table 3.1)
2. Displaying the descriptive statistics for a group of variables (*UE*, Table 3.1)
3. Displaying the simple correlation coefficients between all pairs of variables in a group (*UE*, Table 3.1)
4. Running a simple regression for Woody's Restaurants example (*UE*, Table 3.2)
5. Documenting the results
6. Displaying the actual, fitted, residual, and a plot of the residuals (*UE*, Table 3.2)

The Woody's Restaurants data will be used to show how EViews can be used to do the items listed above. The steps required to display the 'spreadsheet view', descriptive statistics, and simple correlation coefficients between all pairs of variables in a group are as follows:

Displaying the spreadsheet view of a group of variables (*UE*, Table 3.1):

Step 1. Open the EViews workfile named *Woody3.wf1* by clicking **File/Open/Workfile** on the main menu and double click on the file named *Woody3.wf1*.
Step 2. To Create an EViews group for the Woody's Restaurants example, hold down the **Ctrl** button, click on *Y, N, P & I*, select **Show** from the workfile toolbar, and click **OK**.
Step 3. When you click **OK**, EViews displays the spreadsheet view of the data for the variables included in the group (i.e., *Y, N, P & I* in this case). EViews allows you to change the view with the click of a button. If you are in another view, you can display the spreadsheet view by clicking **View/Spreadsheet** on the group window menu bar.
Step 4. Click **Name** on the group menu bar and enter *GROUP01* in the Name to identify object: window.[1]

Displaying the descriptive statistics for a group of variables (*UE*, Table 3.1):

Step 1. Open the group object created in Display the spreadsheet view of a group of variables by double clicking the ▣ group01 icon in the workfile window.
Step 2. Click **View/Descriptive Stats/Individual Samples** on the group window menu bar to view descriptive statistics for *GROUP 01*.[2]

---

[1] You must name a group object if you want to keep its results. Unnamed objects are labeled "UNTITLED" and the results are lost when the object window or the workfile is closed. To name a group, click **Name** on the group menu bar and enter the name in the Name to identify object: window. Once named, a 'group object' is saved with the workfile and can be viewed by double clicking its icon (i.e., ▣ group01) in the workfile window.

[2] The group view drop-down menu is divided into four blocks:
· The views in the first block provide various ways of looking at the actual data in the group.
· The views in the second block display various basics statistics.
· The views in the third block are for specialized statistics for time series data.
· The fourth block contains the label view, which provides information regarding the group object.

The descriptive statistics printed in the table below are defined in footnote 3 at the bottom of this page.[3] See Chapter 16 to gain a better understanding of these statistics.

Step 3. Select **View/Spreadsheet** on the group window menu bar to restore the spreadsheet view for *GROUP 01*.

	Y	N	P	I
Mean	125634.6	4.393939	103887.5	20552.58
Median	122015.0	4.000000	95120.00	19200.00
Maximum	166755.0	9.000000	233844.0	33242.00
Minimum	91259.00	2.000000	37852.00	13240.00
Std. Dev.	22404.09	1.919300	55884.51	5141.865
Skewness	0.355246	0.555101	0.672915	0.933694
Kurtosis	1.920334	2.359612	2.280488	3.161758
Jarque-Bera	2.296908	2.258639	3.202315	4.830791
Probability	0.317127	0.323253	0.201663	0.089332
Observations	33	33	33	33

Displaying the simple correlation coefficients between all pairs of variables in a group (*UE*, Table 3.1):

Step 1. Open the group object created in Display the spreadsheet view of a group of variables by double clicking the **G** group01 icon in the workfile window.

Step 2. Click **View/Correlations** on the group window menu bar to display the simple correlation coefficients between all pairs of variables included in the group object (see the table below).[4]

	Y	N	P	I
Y	1.000000	-0.144225	0.392568	0.537022
N	-0.144225	1.000000	0.726251	-0.031534
P	0.392568	0.726251	1.000000	0.245198
I	0.537022	-0.031534	0.245198	1.000000

---

[3] All of the statistics are calculated using observations in the current sample.
 Mean is the average value of the series, obtained by adding up the series and dividing by the number of observations. Median is the middle value (or average of the two middle values) of the series when the values are ordered from the smallest to the largest. The median is a robust measure of the center of the distribution that is less sensitive to outliers than the mean. Max and Min are the maximum and minimum values of the series in the current sample. Std. Dev. (standard deviation) is a measure of dispersion or spread in the series. Skewness is a measure of asymmetry of the distribution of the series around its mean. Kurtosis measures the peakedness or flatness of the distribution of the series. Jarque-Bera is a test statistic for testing whether the series is normally distributed. The test statistic measures the difference of the skewness and kurtosis of the series with those from the normal distribution. Under the null hypothesis of a normal distribution, the Jarque-Bera statistic is distributed as with 2 degrees of freedom. The reported Probability is the probability that a Jarque-Bera statistic exceeds (in absolute value) the observed value under the null—a small probability value leads to the rejection of the null hypothesis of a normal distribution.
[4] **View/Correlations** displays the correlation matrix of the series in the selected group. Observations for which any one of the series has missing data are excluded from the calculation.

Running a simple regression for Woody's Restaurants example (*UE* 3.2):

To regress the number of customers served (*Y*) on the Constant (*C*), the number of direct market competitors within a two-mile radius of the Woody's location (*N*), the number of people living within a three-mile radius of the Woody's location (*P*), and the average household income of people living within a three-mile radius of the Woody's location (*I*):

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To estimate *UE*, Equation 3.6, p. 74, select **Objects/New Object/Equation** on the workfile menu bar.[5] You can name the equation object now by deleting *Untitled* in the Name for Object: window and typing a name for the equation you are about to estimate, or you can skip this step and name the equation later (if you find that it is worth saving). Click **OK** to reveal the Equation Specification: window.
Step 3. Enter *Y C N P I* in the Equation Specification: window.
Step 4. Do not change the default settings for Method: and Sample:.
Step 5. Click **OK** to get the computer output printed in the top box of *UE*, Table 3.2, p. 76.
Step 6. To name the equation for later use, select **Name** on the equation window menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.[6]

Documenting the results:

All of the data necessary to write the results printed in (*UE*, Equation 3.7, p. 77) are printed in the EViews regression output. EViews provides a variety of views for regression results. For example, if you are documenting the results in a word processing file, you can facilitate the process by clicking **View/Representations** on the equation window menu bar to get the following:

Estimation Command:
=====================
LS Y C N P I

Estimation Equation:
=====================
Y = C(1) + C(2)*N + C(3)*P + C(4)*I

Substituted Coefficients:
=====================
Y = 102192.4277 - 9074.674399*N + 0.3546683674*P + 1.287923391*I

You can copy the last line and paste it into your word processing file to get the first line of *UE*, Equation 3.7.

---

[5] Alternately, select **Quick/Estimate Equation** from the main menu. If this method is used, you must name the equation to save it. Click **Name** on the equation menu bar and enter the desired name in the Name for object: window and click **OK**.
[6] Restrictions on naming an object include, the name cannot exceed 16 characters and it cannot be a reserved name. The following names are reserved and should not be used for series: ABS, ACOS, AR, ASIN, C, CON, CNORM, COEF, COS, D, DLOG, DNORM, ELSE, ENDIF, EXP, LOG, LOGIT, LPT1, LPT2, MA, NA, NRND, PDL, RESID, RND, SAR, SIN, SMA, SQR, and THEN.

Select **View/Estimation Output** on the group window menu bar to restore the estimation output view for *EQ01*.

Displaying the actual, fitted, residual, and a plot of the residuals for a regression (see *UE*, Table 3.2, p. 76):

Step 1. Open the EViews workfile named *Woody3.wf1* and open the equation named *EQ01* by double clicking the equation icon in the workfile window.

Step 2a. Click **View/Actual,Fitted,Residual/Actual,Fitted,Residual Table** on the equation window menu bar (see figure below left). Click **OK** to reveal the figure below right.
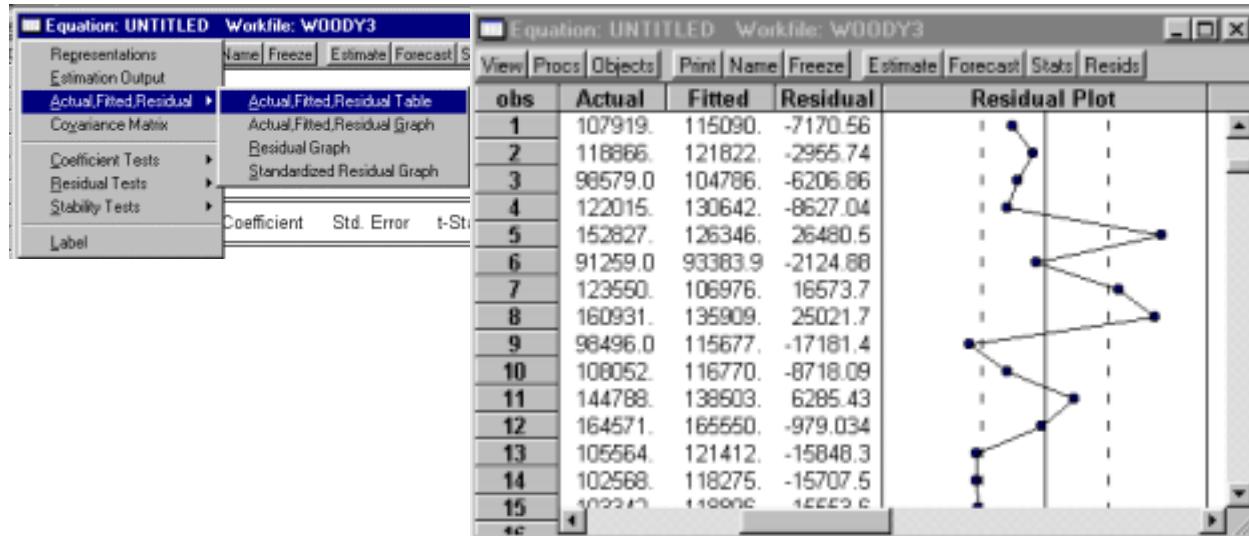
Step 2b. Alternately, to display a graph of the actual, fitted, and residuals for a regression, click **View/Actual,Fitted,Residual/Actual,Fitted,Residual Graph** (see figure below left). Click **OK** to reveal the figure below right. Other views available in the equation window will be explained in future chapters of this guide.
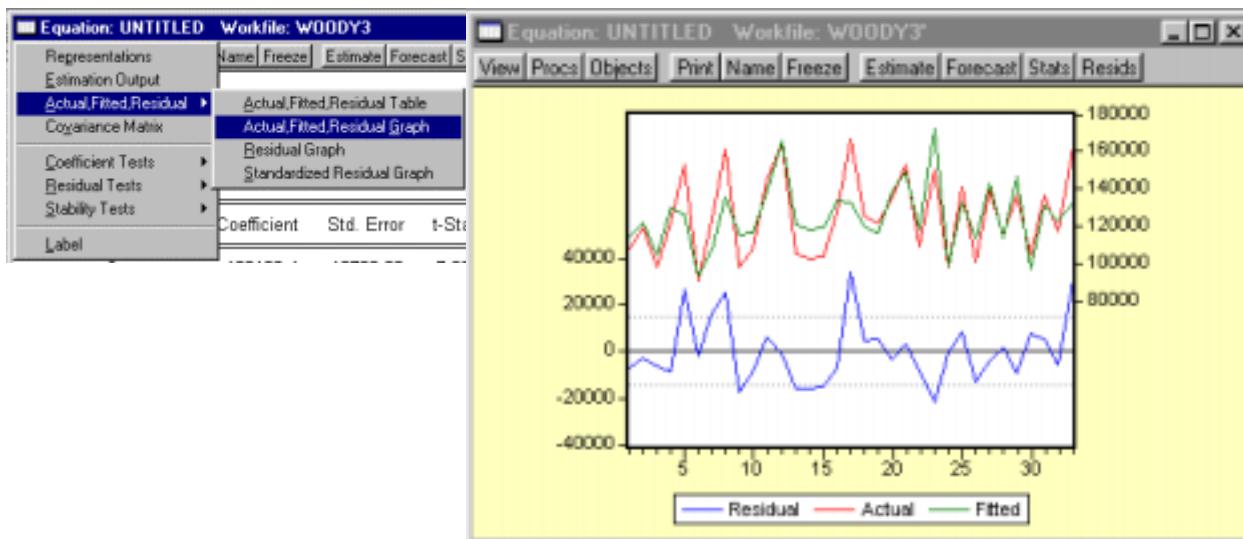
Chapter 4: The Classical Model (OPTIONAL)

> **Read this first:** The procedures described in this chapter are not essential to your understanding of econometrics or to using EViews to reproduce the results published in *UE*. The purpose of describing the procedures in this chapter is to familiarize you with the Monte Carlo Simulation process described in (*UE* 4.3.2, p. 102).

Demonstrate that the estimated βs are drawn from a normal distribution (*UE* 4.3.2, pp. 101-102):

> Follow these steps to complete the process described in (*UE* 4.3.2):

Step 1. Open EViews and type the following commands in the command window, hitting **Enter** after each command:

> CREATE MONTECARLO U 1 15
> MATRIX(20,1) BETA
> SERIES X=10+NRND

> **What does this do?** The first command creates an undated workfile named *MONTECARLO* with 15 observations, the second command creates a matrix named *BETA* with 20 rows[1] and one column to store the sample βs, and the third command creates a series named *X* equal to 10 plus a random number drawn from a normal distribution with zero mean and unit variance.

Step 2. Type the following commands in the command window, press **Enter** after each command:

> SERIES Y=X+0.25*@RNORM
> EQUATION EQ1.LS Y X
> BETA(1)=@COEFS(1)

> **What this does:** The first command generates a random number named *y* (*UE*, Equation 4.11, p. 101), the second command estimates the regression with *y* as the dependent variable and *x* as the independent variable (no constant), and the third command saves the β coefficient on *x* in the first row of the matrix named *BETA*. You have to repeat **Step 2** for each new sample β. However, you don't have to re-type each line, each time. Just put the cursor on the command line and press **Enter**. Change the number (in parenthesis after beta in the command BETA(1)=@COEFS(1)) to the next higher number so you don't overwrite the previous sample β. The next iteration should look like this:

> SERIES Y=X+0.25*@RNORM
> EQUATION EQ1.LS Y X
> BETA(2)=@COEFS(1)

---

[1] Enter the number that equals the number of sample βs you plan to estimate. The number 20 was selected for this example, but for a realistic Monte Carlo simulation, this number should be much larger.

Step 3. Type the following commands in the command window, press **Enter** after each command:

BETA.WRITE(T=XLS) EXCEL
CREATE BETAWF U 1 20
READ(T=XLS) EXCEL 1
RENAME SER01 BETA
BETA.HIST
SAVE

The first command writes the matrix named *BETA* as a series to an Excel file named *EXCEL*, the second command creates a new undated EViews workfile named *BETAWF* with 20 observations,[2] the third command reads the series named *excel* into the EViews workfile and names it *SER01*, the fourth command renames *SER01* to *BETA*, the fifth command creates a histogram of the sample βs, and the last command saves the EViews file.

---

[2] This number should be set equal to the number of rows in the matrix named *beta*. See footnote 1.

The three figures below show how the probability distribution of the estimated sample β series more closely approximates the normal distribution as the number of observations increase from 20 to 2,000 and finally to 8,000. For an explanation of the histogram and descriptive statistics represented, see Chapter 16.

The graphic on the right shows the **View/Descriptive Statistics/Histogram and Stats** EViews for beta with 20 observations.
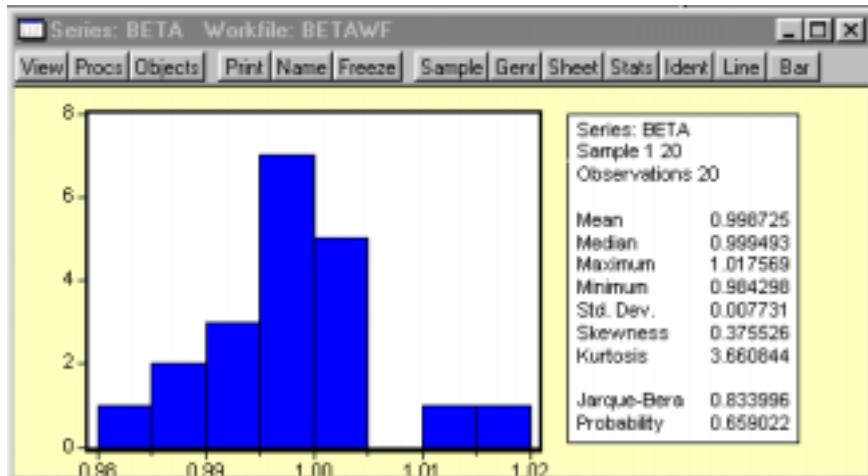
The graphic on the right shows the **View/Descriptive Statistics/Histogram and Stats** EViews for beta with 2,000 observations.
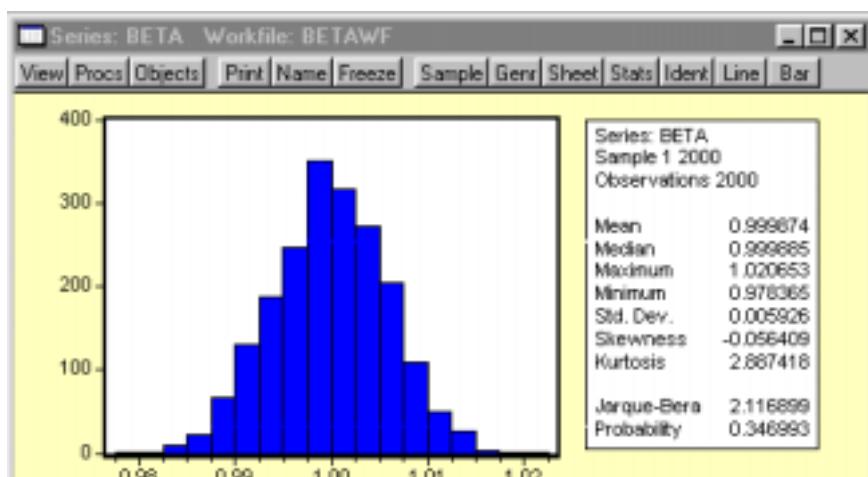
The graphic on the right shows the **View/Descriptive Statistics/Histogram and Stats** EViews for beta with 8,000 observations.
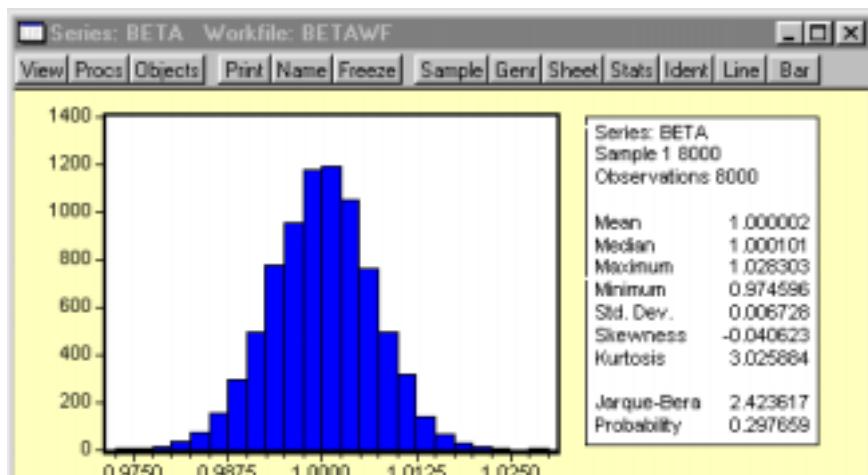
Chapter 5: Basic Statistics and Hypothesis Testing

In this chapter:

Viewing the t-value from an OLS regression (*UE* 5.2.1):

The Woody's Restaurants example is used to explain the use of t-values to test hypotheses concerning the coefficients on the independent variables in an OLS regression model. Follow these steps to open the Woody's Restaurant workfile in EViews and run the regression for the equation $Y_t = \hat{\beta}_0 + \hat{\beta}_N N_t + \hat{\beta}_p P_t + \hat{\beta}_i I_t + e_t$.

Step 1. Open the EViews workfile named *Woody3.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C N P I* in the Equation Specification: window, and click **OK.** EViews generates the following output (also printed in *UE*, Table 3.2, p. 76):

| Dependent Variable: Y				
Method: Least Squares				
Date: 05/23/00 Time: 05:55				
Sample: 1 33				
Included observations: 33				
Variable	Coefficient	Std. Error	t-Statistic[1]	Prob.
C	102192.4	12799.83	7.983891	0.0000
N	-9074.674	2052.674	-4.420904	0.0001
P	0.354668	0.072681	4.879810	0.0000
I	1.287923	0.543294	2.370584	0.0246
R-squared	0.618154	Mean dependent var		125634.6
Adjusted R-squared	0.578653	S.D. dependent var		22404.09
S.E. of regression	14542.78	Akaike info criterion		22.12079
Sum squared resid	6.13E+09	Schwarz criterion		22.30218
Log likelihood	-360.9930	F-statistic		15.64894
Durbin-Watson stat	1.758193	Prob(F-statistic)		0.000003

All information needed for hypothesis testing using the t-test is found in the middle of the EViews equation output table (highlighted in yellow). The first column identifies the name of the variable. The second column reports the estimated coefficient ($\hat{\beta}_k$) for each variable, and the third column reports the standard error for the estimated coefficient (SE $\hat{\beta}_k$). The fourth column prints the calculated t-value given that the border value implied by the null hypothesis (β_{Ho}) is zero (i.e., the t-value in this case is ($\hat{\beta}_k$)/(SE $\hat{\beta}_k$)).

---

[1] The EViews program uses the term t-Statistic rather than the term t-value, which is used in *UE* and this guide.

Calculating critical t-values and applying the decision rule (*UE 5.2.2*):

The critical t-value (t_c) is the value that separates the "acceptance" region from the "rejection" region. Look up this value in *UE*, Statistical Table B-1, p.607. Its value depends on the degrees of freedom (printed in column one of Table B-1) and the level of Type I error specified (i.e., the number at the top of the column in Table B-1 for two-tailed tests or double the value for one-tailed tests). Alternately, you can follow these steps to have EViews calculate the one-tailed and two-tailed, 5% significance level critical t-values (t_c):

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C N P I* 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 vector object named *result* with 10 rows (to store the results of the test statistics for Woody's Restaurants regression), type the following command in the command window: *vector(10) result* and press **Enter**.[2]
Step 5. To compute the two-tailed critical t-value (t_c) for the 5% significance level and save the value in the first row of the vector object named *result*, type the following equation in the command window and press **Enter**:[3]
result(1)=@qtdist(.975,(eq01.@regobs-eq01.@ncoef)).[4]
Step 6. To compute the one-tailed critical t-value (t_c) for the 5% significance level and save the value in the second row of the vector object named *result*, type the following equation in the command window and press **Enter**:
result(2)=@qtdist(.95,(eq01.@regobs-eq01.@ncoef)).
Step 7. Double click on the Vector Object named *result* in the workfile window to view the two-tailed and one-tailed, 5% significance level critical t-value (t_c) for the Woody's Restaurants regression. In this case, the value reported in row one of the *result* vector is *2.045230* and the value reported in row two is *1.69912702653*.

Population was hypothesized to have a positive effect on the number of customers eating at Woody's Restaurants. This implies that the coefficient on *P* is expected to be positive and a one-tailed test is appropriate. Thus, the hypothesis that the coefficient is zero ($\hat{\beta}_p = 0$) is rejected at the 5% significance level because the calculated t-value (*4.879810*) is greater than the one-tailed critical t-value (5% level of significance), calculated in **Step 6** to be *1.69912702653*.

While it is important to be able to apply the decision rule by comparing the calculated t-value (reported in the EViews output) with the critical t-value just calculated, EViews makes it

---

[2] Alternately, select **Objects/New Object/Matrix-Vector-Coef** from the main menu or the workfile menu. Write *result* in the Name for Object: window, and click **OK**. Select *Vecto* in the Type window, enter *10* in the Rows window, enter *1* in the Columns window, and click **OK**.
[3] The command, *@qtdist(p,v)*, calculates the value where the cumulative density function (CDF) of the t-distribution with (*v*) degrees of freedom equals (*p*) probability, leaving *1-p%* of the t-distribution in each tail. Note that you should enter *0.975* for "*p*" for the two-tailed 5% significance level calculation and *0.95* for "*p*" for the one-tailed 5% significance level. In this case, *v* equals (*eq01.@regobs-eq01.@ncoef*), which calculates the degrees of freedom for equation *eq01*. The {*eq01.*} Part can be omitted if the calculation relates to the last regression run.
[4] If you omit the *result(1)*, EViews returns a scalar value on the status line located in the lower left of your screen.

possible to test the null hypothesis that a coefficient is zero (i.e., $\beta_{Ho} = 0$) without knowing the critical t-value. Instead, you can examine the probability (Prob.) value in the last column of the EViews OLS regression output (see the *EQ01* Estimation Output table). The Prob. value shows the probability of drawing a t-value as extreme as the one actually observed when, in fact, the coefficient value is actually zero[5]. This probability is also known as the p-value or the marginal significance level. In terms of *UE*, it represents the probability of making a Type I error if the null hypothesis, that the coefficient is zero, is rejected. Given a p-value, you can tell at a glance if the null hypothesis, that the true coefficient is zero against a two-sided alternative that it differs from zero, should be rejected or accepted. For example, a p-value lower than .05 suggests rejection of the null hypothesis, for a two-tailed test at the 5% significance level. The appropriate probability is one-half that reported by EViews, for a one-sided test. This is comparable to reading the value at the top of the Critical t-Value table that is double the significance level that you are testing for a one-tailed test (i.e., using 0.10 instead of 0.05 for 5% significance level).

Applying the Prob. value, one-tailed test to the population (*P*) variable in the Woody's Restaurants regression, the hypothesis that the coefficient is zero ($\hat{\beta}_p = 0$) is rejected at the 5% significance level if one-half of the Prob. value in the last column is less than or equal to 0.05. Note that the null hypothesis is also rejected at the 1% significance level.

Calculating confidence intervals (*UE* 5.2.4):

Complete steps 1-4 of the section entitled Calculating critical t-values and applying the decision rule before attempting this section (i.e., an equation object named *EQ01* and a vector object named *result* with 10 rows should already be present in the workfile). To calculate and record the 90% confidence interval for a coefficient using EViews:

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To calculate the lower value for the 90% confidence interval for the population coefficient, enter the following formula in the command window, and press **Enter**:[6]
*result(3)= eq01.@coefs(3)-(@qtdist(.95,(eq01.@regobs-eq01.@ncoef)))*eq01.@stderrs(3)*.
Step 3. To calculate the upper value for the 90% confidence interval for the population coefficient, enter the following formula in the command window, and press **Enter**:
*result(4)= eq01.@coefs(3)+(@qtdist(.95,(eq01.@regobs-eq01.@ncoef)))*eq01.@stderrs(3)*.
Step 4. To view the lower and upper confidence interval values, double click the vector icon named *result*. Note that the values *0.231175* and *0.478162* are printed in rows three and four respectively (the same values are reported in *UE*, p. 128).

---

[5] Under the assumption that the errors are normally distributed, or that the estimated coefficients are asymptotically normally distributed.

[6] *eq01.@coefs(i)* and *eq01.@stderrs(i)* are scalar values of the coefficient and standard error of the ith variable in regression *eq01,*where i represents the coefficient number (including the constant) listed in the EViews OLS Estimation Output. Since the population (*P*) variable is listed third in the Estimation Output, *@coefs(3)* and *@stderrs(3)*, calculates the value for the population coefficient and its standard error respectively. As in section 5.2.2, *(@qtdist(.95,(eq01.@regobs-eq01.@ncoef)))* calculates the value where the cumulative density function (CDF) of the t-distribution equals *0.95* probability, leaving 5% of the t-distribution in each tail. The term *(eq01.@regobs-eq01.@ncoef)* calculates the degrees of freedom for *EQ01*, with *eq01.@regobs* calculating the number of observations used to estimate *EQ01* and *eq01.@ncoef* calculating the number of coefficients estimated, including the constant.

Performing the t-test of the simple correlation coefficient (*UE* 5.3.3):

Complete steps 1-4 of the section entitled Calculating critical t-values and applying the decision rule before attempting this section (i.e., an equation object named *EQ01* and a vector object named *result* with 10 rows should already be present in the workfile). To use the t-test to determine whether a particular simple correlation coefficient between *Y* and *P* is significant:

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To calculate the simple correlation coefficient (r) and store it in the fifth row of the result vector, type the following command in the command window, and press **Enter**: *result(5)= @cor(y,p).*
Step 3. To convert the simple correlation coefficient between *Y* and *P* into a t-value and store it in row six of the *result* vector, type the following command in the command window, and press **Enter**: *result(6) = (@cor(y,p)*((@obs(y)-2)^.5))/((1-@cor(y , p)^2)^.5).*
Step 4. To calculate the critical t-value (t_c) for the t-distribution with *@obs(y)-2* degrees of freedom and store it in the seventh row of the *result* vector, type the following command in the command window, and press **Enter**: *result(7) = @qtdist(.975,(@obs(y)-2)).*
Step 5. Double click the icon for the *result* vector to view the results in rows 5, 6, and 7.

Performing the F-test of overall significance (*UE* 5.5):

Complete steps 1-4 of the section entitled Calculating critical t-values and applying the decision rule before attempting this section (i.e., an equation object named *EQ01* and a vector object named *result* with 10 rows should already be present in the workfile). The F-statistic is reported as (15.64894) in the *EQ01* estimation output table. Follow these steps to calculate the F-statistic from *EQ01* and save it in row seven of a vector named *result*:

Step 1. Open the EViews workfile named *Woody3.wf1*.
Step 2. To store the F-statistic from *EQ01* in row eight of the results vector, type the following command in the command window, and press **Enter**: *result(8)=eq01.@f.*
In this case, the F-statistic tests the hypothesis that all of the slope coefficients excluding the constant, in *EQ01*, are zero. Under the null hypothesis with normally distributed errors, this statistic has an F-distribution with k degrees of freedom in the numerator and n-k-1 degrees of freedom in the denominator, where n equals the number of observation and k equals the number of independent variables (k does not include the constant) in the model. The null hypothesis can be rejected if the calculated F-statistic exceeds the critical F-value at the chosen level of significance. The critical F-value can be determined from *UE*, Appendix B, Statistical Tables B-2 or B-3, depending on the level of significance chosen. Alternately,
Step 3. To have EViews calculate the 5% critical F-value for *EQ01* and store it in row nine of the results vector, type the following command in the command window, and press **Enter**: *result(9)=@qfdist(0.95,eq01.@ncoef-1,eq01.@regobs- eq01.@ncoef).*
Step 4. Double click the icon for the *result* vector to view the results in rows 8 and 9.

test statistic	result row	value
t-critical for regression - 5% level of significance (two-tailed test) =	R1	2.04523
t-critical for regression - 5% level of significance (one-tailed test) =	R2	1.699127
lower confidence interval =	R3	0.231175
upper confidence interval =	R4	0.478162
The simple correlation coefficient (r) =	R5	0.392568
t-calculated for correlation =	R6	2.376503
t-critical for correlation =	R7	2.039513
The F-statistic =	R8	15.64894
Critical value of the F-statistic - 5% level of significance =	R9	2.93403
	R10	0

Since the 5% critical F-value for *EQ01* (i.e., *2.934030*) is significantly less than the calculated F-statistic (i.e., *15.64894*), we can reject the null hypothesis that all of the slope coefficients in *EQ01* are zero.

The p-value printed just below the F-statistic in the EViews regression output, denoted Prob(F-statistic), represents the marginal significance level of the F-test. If the p-value is less than the significance level you are testing, say .05, you reject the null hypothesis that all slope coefficients are equal to zero. For *EQ01*, the p-value is 0.000003, so we reject the null hypothesis that all of the regression coefficients are zero. Note that the F-test is a joint test so that even if all the t-values are insignificant, the F-statistic can be highly significant.

Exercises:

16. EViews can be used to complete parts 1, b, c & f of exercise 16.
    a. Review the section Calculating critical t-values and applying the decision rule to learn how to use EViews to calculate the critical values for testing your hypothesis concerning the regression coefficients in this exercise.
    b. Review the section Performing the F-test of overall significance to learn how to use EViews to calculate the critical value for the F-test of the overall significance of the estimated equation.
    c. Review the section Calculating confidence intervals to learn how to use EViews to determine lower and upper confidence intervals for an estimated coefficient.
    d.
    e.
    f. Review the section Using EViews to estimate a multiple regression model of beef demand in Chapter 2, if you have trouble estimating this multiple regression model using EViews.

Chapter 6: Specification: Choosing the Independent Variables

Adding or deleting variables to/from an OLS model in EViews (*UE* 6.1.2 - 6.3):

EViews makes it easy to try alternative versions of an OLS model in order to determine whether omitting a variable is likely to result in specification bias or whether the variable is irrelevant. The four important specification criteria (*UE*, pp. 167-168) don't always agree. However, when the theory is not absolutely clear about the relevancy of including a specific variable in a model, the other three criteria (i.e., t-test, adjusted R^2, and bias) should be considered. The only way to check these criteria is to run the regression with and without the variable and evaluate the results in terms of t-test, adjusted R^2, and bias. The following steps outline a procedure to determine whether the price of beef (PB) is a relevant variable in the demand for chicken model (*UE*, Equation 6.8, p. 160):

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* in the Equation Specification: window, and click **OK**.

Step 3. To preserve this EViews Estimation Output view of *UE*, Equation 6.8, p. 160, for later comparison, select **Name** on the equation menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.[1]

Step 4. Create a duplicate copy of *EQ01* by selecting **Objects/Copy object…** on the *EQ01* menu bar. A new UNTITLED copy of *EQ01* Estimation Output appears. In this new equation window, select **Estimate** on the equation menu bar, delete PB from the Equation Specification: window, and click **OK**.

Step 5. To preserve this EViews Estimation Output of *UE*, Equation 6.9, p. 161, for later comparison, select **Name** on the equation menu bar, enter *EQ02* in the Name to identify object: window, and click **OK**.

Step 6. Compare and evaluate the two equations based on t-statistics, adjusted R^2, and bias.

---

[1] Alternately, the EViews Estimation Output could have been preserved by selecting **Freeze** on the equation menu bar. The **Freeze** button on the objects toolbar creates a duplicate of the current view of the original object. The primary feature of freezing an object is that the tables and graphs created by freeze may be edited for presentations or reports. Frozen views do not change when the workfile sample is changed or when the data change. The purpose for freezing the regression output table is to allow us to view it later by double clicking the objects icon in the workfile window. In order to do that, the frozen object must be named.

Lagging variables in an OLS model using EViews (*UE* 6.5):

EViews makes it easy to lag variables in an equation.[2] Equations 6.22 & 6.23 refer to a hypothetical model and they are not actually estimated in *UE*. However, the demand for chicken model (*UE*, Equation 6.8, p. 160) will be used to show how to lag variables in EViews.

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

Step 2. To run the regression for Y_t on PC_{t-1}, PB_t and YD_t, select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C PC(-1) PB YD* in the Equation Specification: window (see the graphic to the right), and click **OK**.
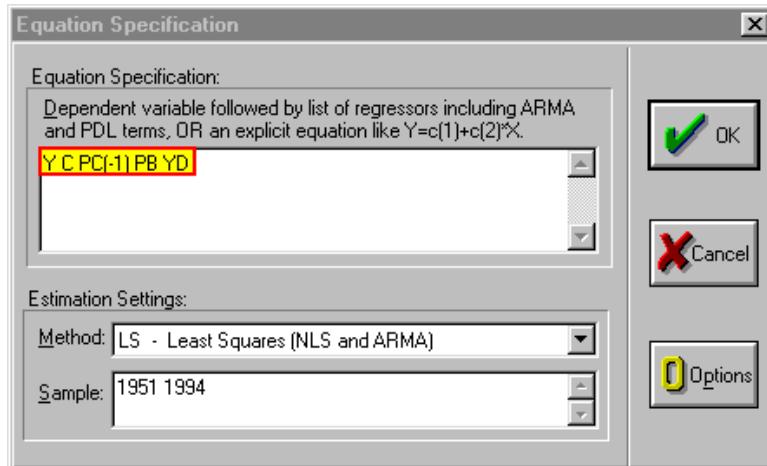
Note that EViews reports that it has adjusted the sample (see the graphic to the right). The range and sample in the workfile window show *1951 1994* but the equation output reports *Sample(adjusted): 1952 1994*. You should be aware that if you include lagged variables in a regression, the degree of sample adjustment will

differ depending on whether data for the pre-sample period are available or not. For example, suppose the workfile range is *1950 1994* and the workfile sample is *1950 1994*. If you specify a regression with PC lagged one period, EViews will not adjust the sample because it can use the data for 1950 in the workfile.

---
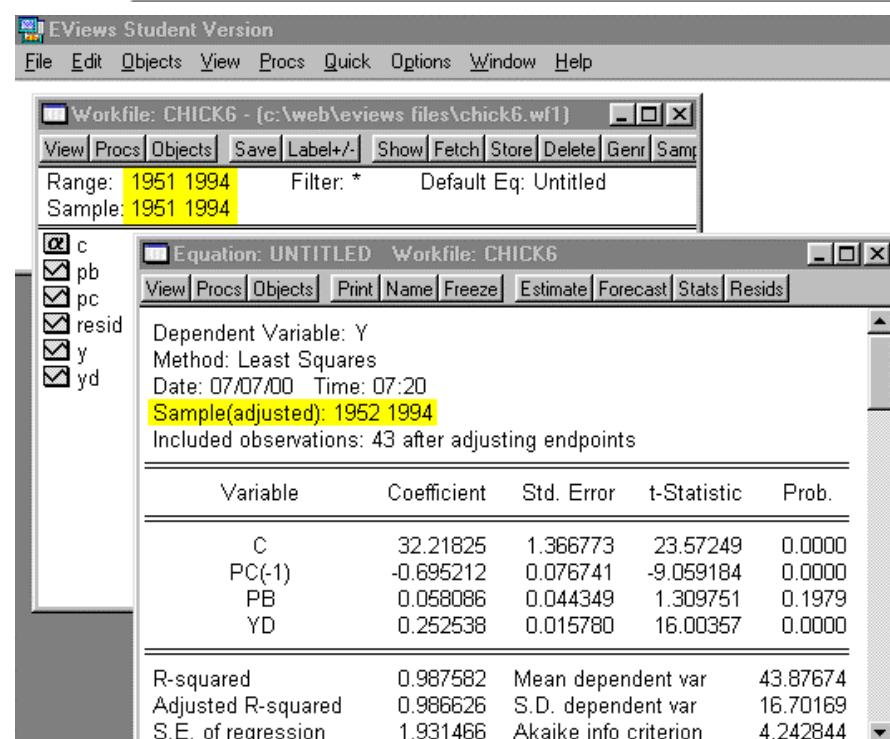
[2] In fact, nearly any transformation of the variables using EViews functions is allowed. See **Help/Reference(Commands and Functions)/Function Reference** for a list of EViews functions.

Ramsey's Regression Specification Error Test (RESET) (*UE* 6.8.1):

Complete **Steps 1 - 5** of the section entitled Adding or deleting variables to/from an OLS model in EViews before attempting this section (i.e., *EQ02* should be present in the workfile). Follow these steps to carry out the Ramsey's Regression Specification Error Test (RESET) using the step-by-step approach followed in *UE*, pp. 193-195:

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

Step 2. Open *EQ02* by double clicking its icon in the workfile window (see *UE*, Equation 6.9, p. 194).

Step 3. Select **Forecast** on the equation menu bar, enter *YF* in the Forecast name: window, and click **OK**.[3]

Step 4. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C PC YD YF^2 YF^3 YF^4* in the Equation Specification: window, and click **OK** (see *UE*, Equation 6.10, p. 194).

Step 5. Select **Name** on the equation menu bar, enter *EQ03* in the Name to identify object: window, and click **OK**.

Step 6. Select **View/Coefficient Tests/Wald-Coefficient Restrictions** on the equation menu bar, enter *C(4)=0, C(5)=0, C(6)=0* in the Coefficient restrictions separated by commas: window,[4] and click **OK** to get the graphic shown on the right.

Step 7. Determine the critical F-statistic from *UE*, Table B-2 or by entering the formula: *=@qfdist(0.95,3,eq03. @regobs- eq03.@ncoef)* in the command window, pressing **Enter**, and reading the value on the status line in the lower left of the screen.

Equation: EQ03 Workfile: CHICK6	_ □ ×
View Procs Objects Print Name Freeze Estimate Forecast Stats Resids	

Wald Test:
Equation: EQ03

Null Hypothesis:	C(4)=0
	C(5)=0
	C(6)=0

| F-statistic | 4.323568 | Probability | 0.010205 |
| Chi-square | 12.97070 | Probability | 0.004700 |

Step 8. The F-statistic, highlighted in yellow, is the same as reported in *UE*, p. 195.[5] Since the calculated F-statistic of 4.32 exceeds the critical F-statistic of 2.85, the null hypothesis that the coefficients on the added variables are jointly zero can be rejected at the 5% level. This is in spite of the fact that none of their coefficients are individually significant.

---
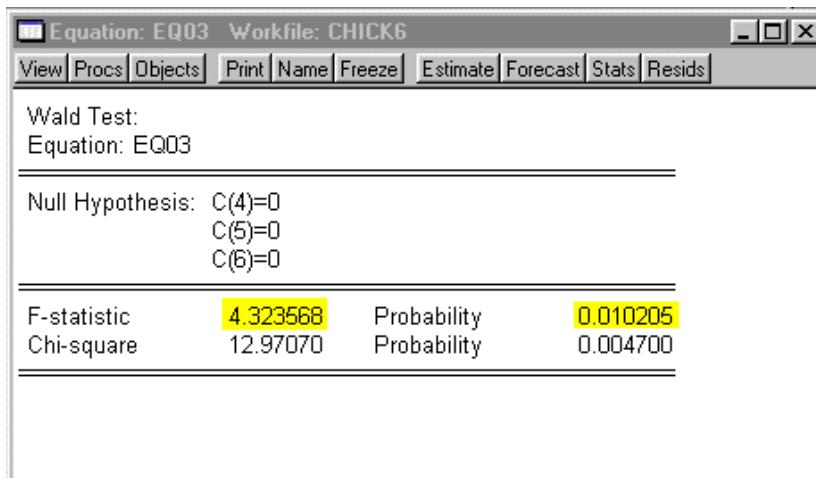
[3] This creates a new series with forecast values of Y based on the estimated coefficients for *EQ02*.

[4] Note that the coefficient restrictions are written as C(i), where i represents the coefficient order number of the variable as it was entered in the Equation Specification: window, following the dependent variable. Thus, *C(4),C(5)* and *C(6)* represent the coefficients for *YF^2 YF^3 YF^4* in the Equation Specification: *Y C PC YD YF^2 YF^3 YF^4*.

[5] The Chi-square statistic is equal to the F-statistic times the number of restrictions under test. In this example, there are three restrictions, so the Chi-square test statistic is three times the size of the F-statistic, but the p-values of both statistics indicate that we can decisively reject the null hypothesis that the three coefficients are zero.

Ramsey's Regression Specification Error Test (RESET) (*EViews*):

Complete **Steps 1 - 5** of the section entitled Adding or deleting variables to/from an OLS model in EViews before attempting this section (i.e., *EQ02* should be present in the workfile). Follow these steps to carry out the Ramsey's Regression Specification Error Test (RESET) using EViews, built in Ramsey's RESET Test:

Step 1. Open the EViews workfile named *Chick6.wf1*.
Step 2. Open *EQ02* by double clicking its icon in the workfile window (see *UE*, Equation 6.9, p. 194).
Step 3. Select **View/Stability Tests/Ramsey RESET Test…**, enter *3* in the Number of fitted terms: window[6], and click **OK** to get the table below.

Note that the output is very similar to the output achieved with the step-by-step approach in the previous section. In this case, the test results are printed above the regression output table.

Since the calculated F-statistic of 4.32 exceeds the critical F-statistic of 2.85, the null hypothesis that the coefficients on the added variables are jointly zero can be rejected at the 5% level. This is in spite of the fact that all of their individual t-statistics are insignificant.

Ramsey RESET Test:				
F-statistic	4.323568	Probability		0.010205
Log likelihood ratio	12.92125	Probability		0.004810
Test Equation:				
Dependent Variable: Y				
Method: Least Squares				
Date: 07/27/00 Time: 07:26				
Sample: 1951 1994				
Included observations: 44				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	23.80305	55.36771	0.429908	0.6697
PC	-0.591937	1.718030	-0.344544	0.7323
YD	0.360179	0.714812	0.503880	0.6173
FITTED^2	0.023868	0.082475	0.289394	0.7739
FITTED^3	-0.000748	0.001106	-0.676301	0.5029
FITTED^4	5.48E-06	5.36E-06	1.022646	0.3129
R-squared	0.988647	Mean dependent var		43.37500
Adjusted R-squared	0.987154	S.D. dependent var		16.83854
S.E. of regression	1.908510	Akaike info criterion		4.256646
Sum squared resid	138.4116	Schwarz criterion		4.499945
Log likelihood	-87.64622	F-statistic		661.8504
Durbin-Watson stat	0.861509	Prob(F-statistic)		0.000000

---

[6] The fitted terms are the powers of the fitted values from the original regression, starting with the square or second power. For example, if you specify 3, then the test will add $\hat{y}^2$, $\hat{y}^3$, and $\hat{y}^4$ in the regression. If you specify a large number of fitted terms, EViews may report a near singular matrix error message since the powers of the fitted values are likely to be highly collinear. The Ramsey RESET test is applicable only to an equation estimated by least squares.

Akaike's Information Criterion (AIC) and the Schwartz Criterion (SC):

Complete **Steps 1 - 5** of the section entitled <u>Adding or deleting variables to/from an OLS model in EViews</u> before attempting this section (i.e., *EQ01* and *EQ02* should be present in the workfile). The Akaike's Information Criterion (AIC) and Schwartz Criterion (SC) are both printed in the <u>Estimation Output</u> of EViews' OLS regressions.

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

Step 2. Open *EQ01* by double clicking its icon in the workfile window (see *UE*, Equation 6.8, p. 160 or 196) to get the <u>Estimation Output</u> below.

Dependent Variable: Y				
Method: Least Squares				
Date: 07/26/00 Time: 09:11				
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. Open *EQ02* by double clicking its icon in the workfile window (see *UE*, Equation 6.9, p. 161) to get the <u>Estimation Output</u> below.

Dependent Variable: Y				
Method: Least Squares				
Date: 07/26/00 Time: 08:01				
Sample: 1951 1994				
Included observations: 44				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	32.94193	1.251191	26.32845	0.0000
PC	-0.700954	0.084099	-8.334841	0.0000
YD	0.272477	0.005936	45.90552	0.0000
R-squared	0.984772	Mean dependent var		43.37500
Adjusted R-squared	0.984030	S.D. dependent var		16.83854
S.E. of regression	2.127957	Akaike info criterion		4.413948
Sum squared resid	185.6562	Schwarz criterion		4.535597
Log likelihood	-94.10685	F-statistic		1325.737
Durbin-Watson stat	0.946570	Prob(F-statistic)		0.000000

Note that the AIC & SC information criterion reported in EViews (see numbers highlighted in yellow) are larger when *PB* is omitted from the OLS regression (i.e., *EQ02*). Both Akaike's and the Schwartz Criterion provide evidence that *UE*, Equation 6.8 (i.e., EViews *EQ01*) is preferable to *UE*, Equation 6.9 (i.e., EViews *EQ02*).

15. Open the EViews workfile named *Drugs.wf1*.

   a.

   i) Select **Objects/New Object/Equation** on the workfile menu bar, enter *P C GDPN CVN PP DPC IPC CV* in the Equation Specification: window, and click **OK**. Select **Name** on the equation window menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.

   ii) Select **Objects/New Object/Equation** on the workfile menu bar, enter *P C GDPN CVN PP DPC IPC N* in the Equation Specification: window, and click **OK**. Select **Name** on the equation window menu bar, enter *EQ02* in the Name to identify object: window, and click **OK**.

   b. Open *EQ01* and *EQ02* at the same time. Use information in these tables, *UE* 6.8 (Appendix) and the procedures outlined in this guide to determine whether *CV* and/or *N* are irrelevant or omitted variables.

   c.

   d.

Chapter 7: Specification: Choosing A Functional Form

In this chapter:
1. Table with EViews specification for functional forms
2. Calculating "Quasi - R^2" in EViews (*UE* 7.3.1, footnote 5, p. 215)
3. Calculating "Quasi - R^2" for a linear versus log-lin model using EViews
4. Coefficient restrictions tests using EViews (*UE*, Appendix 7.7)
5. The *Chow test*, alternately termed Chow's Breakpoint Test (*UE*, Appendix 7.7)

UE section 7.2 presents alternative functional forms that are useful when specifying econometric models. Linear models are frequently too restrictive to properly fit the functional form suggested by the underlying theory.

The last column of Table 7.1 below shows the correct EViews specification for the alternative functional forms printed in *UE*, Table 7.1, p. 214. You can use the table as a guide, but you must realize that Y represents the dependent variable while X_1 & X_2 represent the only independent variables in all of the equations/specifications. Note that a constant (C) should be included in all models even if theory suggests otherwise (see *UE*, p. 201). You must have a workfile open in order to specify and estimate a regression model. Then, to specify a regression model in EViews, select **Objects/New Object/Equation** from the workfile menu and enter the appropriate *EViews specification* (see the last column of the table below), in the Equation Specification: window.[1]

Table 7.1: EViews Specification of Functional Forms

Section	Equation #	Fcn. Form	Equation specification	EViews specification
7.2.1	----	Linear	$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	Y C X$_1$ X$_2$
7.2.2	7.3	Double-Log	$\ln Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2$	log(Y) C log(X$_1$) log(X$_2$)
7.2.3	7.7	Lin-Log	$Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 X_2$	Y C log(X$_1$) X$_2$
7.2.3	7.9	Log-Lin	$\ln Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	log(Y) C X$_1$ X$_2$
7.2.4	7.10	Polynomial	$Y = \beta_0 + \beta_1 X_1 + \beta_2 (X_1^2) + \beta_3 X_2$	Y C X$_1$ X$_1$^2 X$_2$
7.2.5	7.13	Inverse	$Y = \beta_0 + \beta_1 (1/X_1) + \beta_2 X_2$	Y C 1/X$_1$ X$_2$
7.5	7.20	Dummy*	$Y = \beta_0 + \beta_1 X_1 + \beta_2 D_1$	Y C X$_1$ D$_1$
7.5	7.22	Dummy**	$Y = \beta_0 + \beta_1 X_1 + \beta_2 D_1 + \beta_3 D_1 X_1$	Y C X$_1$ D$_1$ D1*X1

* Intercept dummy variable. ** Intercept and slope dummy variables.

Calculating "Quasi - R^2" in EViews (*UE* 7.3.1, footnote 5, p. 215):

The dependent variable must be in the same form when using R^2 and adjusted R^2 to compare the overall goodness of fit between two equations. For example, it would not be appropriate to compare the R^2 for a linear model with a double-log or a log-lin model. However, it would be appropriate to compare R^2 for a linear model with a lin-log, a

---

[1] Alternately, select **Quick/Estimate Equation** from the main menu. If this method is used you must name the equation to save it. Select **Name** on the equation menu bar and enter the desired name in the Name to identify object: window, and click **OK**.

polynomial, or an inverse functional form model. Likewise, it would be appropriate to compare R^2 for double-log and log-lin functional form models. In order to demonstrate the process, the car acceleration data introduced in *UE*, Exercise 16, p. 234, will be used to demonstrate the process of calculating the quasi-R^2. The steps below show how to compare the goodness of fit for models using S (the number of seconds it takes a car to accelerate from 0 to 60 miles per hour) as the dependent variable versus using the natural log of S as the dependent variable. In both models, the independent variables are the same as the original model printed at the top of *UE*, p. 236.

Calculating "Quasi - R^2" for a linear versus a log-lin model using EViews:

Step 1. Open the EViews workfile named *Cars7.wk1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *S C T E P H* in the Equation Specification: window, and click **OK**.

Step 3. Select **Name** on the equation menu bar, write *linear* in the Name to identify object: window, and click **OK**. Minimize the equation object named *linear*.

Step 4. Select **Objects/New Object/Equation** on the workfile menu bar, enter *log(S) C T E P H* in the Equation Specification: window (i.e., the log-lin functional form), and click **OK**.

Step 5. Select **Name** on the equation menu bar, write *loglin* in the Name to identify object: window, and click **OK**.

Step 6. Select **Forecast** on the equation menu bar, select *S* in the Forecast of:[2] window, enter *SF* in the Forecast name: window, uncheck the two boxes in the Output: window (the only objective here is to create a forecast series, not a forecast evaluation), and click **OK**. A new series named *SF* appears in the workfile window.

Steps 7, 8 & 9 calculate the quasi-R^2 for this regression (*UE* 7.3.1, footnote 5, p. 215).

Step 7. Minimize the equation window, select **Genr** on the workfile menu bar, type *numerator=(S-SF)^2* in the Enter equation: window, and click **OK** (this step generates the un-summed variable in the numerator of the quasi-R^2 equation).

Step 8. Select **Genr** on the workfile menu bar, type *denominator=(S-@mean(S))^2* in the Enter equation: window, and click **OK** (this step generates the un-summed variable in the denominator of the quasi-R^2 equation).

Step 9. To calculate the quasi-R^2, type the following equation in the command window and press **Enter**: *scalar quasir2=1-(@sum(numerator)/@sum(denominator))*. A new variable named *quasir2* will appear in the workfile window. Double click on it and the value for the quasi-R^2 will be displayed in the lower left of the screen (0.783958974). The quasi-R^2 calculated in **Step 9** (i.e., 0.78) is in-between the R^2 from the linear model estimated in **Step 2** (i.e., 0.71) and the R^2 from the log-lin model estimated in **Step 5** (i.e., 0.81).

---

[2] The **Forecast** procedure in EViews gives you the option of forecasting the transformed dependent variable (i.e., *LOG(S)* in this case) or the original variable (i.e., *S* in this case). Select S, since the computation of quasi-R^2 requires converting of *LOG(S)* to *S* by taking the anti-log of the dependent variable (this can also be done by using the EViews command *@exp(LOG(S))*.

The F-test can be used to test a wide range of hypothesis concerning regression coefficients. For example, suppose that the claim was made that when a car has a manual transmission it increases its acceleration speed (i.e., decreases the number of seconds it takes to accelerate from 0 to 60 miles per hour) just as much as adding 100 horsepower to the car. Translating this into the language of *UE*, Equation 7.28, p. 235, this means that the absolute value of the coefficient on T_i is 100 times larger than the absolute value of the coefficient on H_i. Just looking at the size of the estimated coefficients, it appears that you can easily reject the hypothesis because the absolute value of the coefficient on T_i is only about 41.5 times larger than the absolute value of the coefficient on H_i (divide the coefficient on T_i by the coefficient on H_i). However, these coefficients are just estimates. Follow these steps to carry out an F-test for the null hypothesis that the absolute value of the coefficient on T_i is 100 times larger than the absolute value of the coefficient on H_i. :

Step 1. Open the EViews workfile named *Cars7.wk1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *S C T E P H* in the Equation Specification: window, and click **OK**.
Step 3. Select **Name** on the equation menu bar, write *EQ01* in the Name to identify object: window, and click **OK**.
Step 4. Select **View/Coefficients Tests/Wald-Coefficient Restrictions …** on the equation menu bar, enter *-C(2)=-100*C(5)* in the Coefficients separated by commas: window, and click **OK** to reveal the following output:[3]

Wald Test:
Equation: EQ01

Null Hypothesis:	-C(2)=-100*C(5)		
F-statistic	2.485049	Probability	0.124472
Chi-square	2.485049	Probability	0.114933

The null hypothesis is *-C(2)=-100*C(5)*, since variable T is the second coefficient and variable H is the fifth coefficient in the EViews Estimation Output from **Step 2**. The F-statistic compares the residual sum of squares computed with and without the restrictions imposed. If the restrictions are valid, there should be little difference in the two residual sum-of-squares and the F-value should be small. Based on the Wald Test: results table, the null hypothesis cannot be rejected at the 5% level of significance. The calculated F-statistic of *2.49* is less than the critical F-value of *4.14*. The critical F-value can be found in *UE*, Table B-2, p. 609 for 1 degree of freedom in the numerator and 33 (interpolate between the 30 and 40) degrees of freedom in the denominator or EViews can calculate

---

[3] The coefficients should be referred to as C(1), C(2), and so on (do not use series names). Multiple coefficient restrictions must be separated by commas and the restrictions should be expressed as equations involving estimated coefficients and constants. The coefficients should be referred to as C(1), C(2), and so on (do not use series names).

its value.[4] The reported probability is the marginal significance level of the F-test. It supports this result in that rejecting the null hypothesis would be wrong less than 12.44% of the time.

The Chi-square statistic is equal to the F-statistic times the number of restrictions under test. In this example, there is only one restriction and so the two test statistics are identical with the p-values of both statistics indicating that we cannot reject the null hypothesis, that the absolute value of the coefficient on T_i is 100 times larger than the absolute value of the coefficient on H_i, at the 10% significance level. The 10% significance critical value for the χ^2 test can be found in *UE*, Table B-8, p. 619 to be 2.71.

The *Chow test*, alternately termed Chow's Breakpoint Test (*UE*, Appendix 7.7):

Chow's Breakpoint Test divides the data into two sub-samples.[5] It then estimates the same equation for each sub-sample separately, to see whether there are significant differences in the estimated equations. A significant difference indicates a structural change in the relationship.

Follow these steps to apply the Chow breakpoint test, as described in *UE*, pp. 241-242, to determine whether there was a structural change in the demand for chicken in 1976:

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* in the Equation Specification: window, and click **OK**.
Step 3. Select **Name** on the equation menu bar, write *EQ01* in the Name to identify object: window, and click **OK**.
Step 4. Select **View/Stability Tests/Chow Breakpoint Test…** on the equation menu bar, enter *1976* in the Enter one date (observation) for the Forecast Test or one or more dates for the Breakpoint Test: window, and click **OK** to reveal the following output:

Chow Breakpoint Test: 1976			
F-statistic	4.542962	Probability	0.004498
Log likelihood ratio	17.98027	Probability	0.001245

EViews reports two test statistics for the Chow breakpoint test. The F-statistic is based on the comparison of the restricted and unrestricted sum of squared residuals. EViews calculates the F-statistic using the formula printed in *UE*, Equation 7.36, p. 242. In this

---

[4] To have EViews calculate the 5% critical F-value for this problem, type the following equation in the command window =@qfdist(0.95,1,eq01.@regobs-eq01.@ncoefs), press **Enter** and view the following value on the status bar in the lower left of the screen [Scalar = 4.13925249556]. For the 10% critical F-value type =@qfdist(0.90,1,eq01.@regobs-eq01.@ncoefs) in the command window, and press **Enter** and view the following value on the status bar in the lower left of the screen [Scalar = 2.86408341558].

[5] One major drawback of the breakpoint test is that each sub-sample requires at least as many observations as the number of estimated parameters. This may be a problem if, for example, you want to test for structural change between wartime and peacetime where there are only a few observations in the wartime sample.

case, the calculated F-statistic of 4.54 exceeds the critical F-value of 2.63 for the 5% level of significance so the null hypothesis of no structural change can be rejected. The critical F-value can be found in *UE*, Table B-2, p. 609 for 4 degrees of freedom in the numerator and 36 (interpolate between the 30 and 40) degrees of freedom in the denominator or EViews can calculate its value.[6] The reported probability is the marginal significance level of the F-test. It supports this result in that rejecting the null hypothesis would be wrong less than 0.4498% of the time.

The log likelihood ratio statistic is based on the comparison of the restricted and unrestricted maximum of the log likelihood function. The LR test statistic has an asymptotic χ^2 distribution with degrees of freedom equal to (m-1)*(k+1) under the null hypothesis of no structural change, where m is the number of sub-samples and k is the number of independent variables in the model (i.e., m = 2 in this case because one breakpoint is selected and k = 3). The calculated value for LR test statistic of 17.98 exceeds of 9.49 for the 5% level of significance and 13.28 for the 1% level of significance so the null hypothesis of no structural change can be rejected.[7] The reported probability is the marginal significance level of the χ^2 test. It supports this result in that rejecting the null hypothesis would be wrong less than 0.1245% of the time.

---

[6] To have EViews calculate the 5% critical F-value for this problem, type the following equation in the command window $=@qfdist(0.95,eq01.@ncoef,eq01.@regobs-2*eq01.@ncoef)$, press **Enter** and view the following value on the status bar in the lower left of the screen Scalar = 2.63353209421.
[7] The critical value for the χ^2 test can be found in *UE*, Table B-8, p. 619.
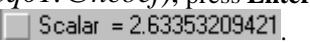
Chapter 8: Multicollinearity

In this chapter:

Perfect multicollinearity (*UE* 8.1.1):

EViews is incapable of generating estimates of regression coefficients when the model specification contains two or more variables that are perfectly collinear. When the equation specification contains two or more perfectly collinear (or even some highly collinear) variables, EViews will put out the error message "*Near singular matrix*."

The next two sections explain how EViews can be used to detect severe multicollinearity (*UE* 8.3). The data for the fish/Pope example found in *UE*, Table 8.1, p. 267, will be used to demonstrate the two methods discussed in *UE*.

Detecting multicollinearity with simple correlation coefficients (*UE* 8.3.1):

High simple correlation coefficients between variables is a sign of multicollinearity. Follow these steps to compute the simple correlation coefficient between variables:

Step 1. Open the EViews workfile named *Fish8.wk1*.
Step 2. Create a group object for the variables found in *UE*, Equation 8.24, *p.* 268 (i.e., *F PF PB log(YD) N P)*. An easy way to create a group object for a set of variables from a regression model is to select **Procs/Make Regressor Group** on the equation window menu bar (refer to Chapter 3 to review the usual way of creating a group object).
Step 3. Select **View/Correlations** on the group object menu bar to reveal the simple correlation coefficients between all of the variables in the group.
Step 4. Select **Freeze** on the group object menu bar to create a table of the simple correlation coefficients. Select **Name** on the table object menu bar to name the table.

	F	*PF*	*PB*	*LOG(YD)*	*N*	*P*
F	1.000000	0.847590	0.818532	0.780012	0.736549	0.585630
PF	0.847590	1.000000	0.958096	0.915320	0.883207	0.734643
PB	0.818532	0.958096	1.000000	0.814890	0.781400	0.663162
LOG(YD)	0.780012	0.915320	0.814890	1.000000	0.945766	0.744500
N	0.736549	0.883207	0.781400	0.945766	1.000000	0.571129
P	0.585630	0.734643	0.663162	0.744500	0.571129	1.000000

To test whether a particular simple correlation coefficient between variables is significant, refer to Performing the t-test of the simple correlation coefficient (*UE* 5.3.3).

Calculating Variance Inflation Factors (*UE 8.3.2*):

The procedure for calculating the Variance Inflation Factors (VIF's) for explanatory variables in a regression model is explained in *UE*, p. 257. Use the following steps to calculate the VIF for the *PF* explanatory variable in *UE*, Equation 8.24, p. 268:

Step 1. Open the EViews workfile named *Fish8.wk1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *PF C PB log(YD) N P* in the Equation Specification: window, and click **OK**. Note the $R^2 = 0.976680$.

Step 3. To save this regression, select **Name** on the equation window menu bar, enter *EQPF* in the Name to identify object: window, and click **OK**.

Step 4. To calculate the VIF for the variable *PF*, write the equation: *scalar VIFPF=1/(1-EQPF.@R2)* in the command window, and press **Enter**. The status line in the lower left of the screen will display the phrase *"VIFPF successfully created."* A new variable named *VIFPF* will appear in the workfile window.

Step 5. Double click on the *VIFPF* icon and the value for the Variance Inflation Factor for the variable named *VIFPF = 42.88...* will be displayed on the status line (lower left of the screen). The fact that *VIFPF* is significantly greater than 5 (i.e., the common rule of thumb value stated in *UE*, p. 258), would suggest that *PF* exhibits severe multicollinearity.

Step 6. Repeat **Steps 2-5** to compute the Variance Inflation Factors for the rest of the explanatory variables in *UE*, Equation 8.24, p. 268.

Transforming multicollinear variables (*UE 8.4.3*):

Variable transformations can be achieved by creating a new variable or by simply writing the transformation in the Equation Specification: window in EViews. In many ways, the latter is preferred because the equation output labels depict the transformation. Otherwise, it is easy to forget what transformations have been made. The table below lists the most common transformations used to rid a model of multicollinearity, along with the EViews specification (***EViews** specification*) to make it happen.

Function Name	*Description*	***EViews** specification*
Linear combination	$Y_t = \beta_0 + \beta_1(X_t + Z_t)$	Y C X+Z*
First difference	$Y_t = \beta_0 + \beta_1(X_t - X_{t-1})$	Y C d(X)
First difference of the logarithm	$Y_t = \beta_0 + \beta_1(\ln X_t - \ln X_{t-1})$	Y C dlog(X)
One-period % change (in decimal)	$Y_t = \beta_0 + \beta_1[(X_t - X_{t-1})/X_t]$	Y C pch(X)

* no spaces in the X+Z variable designation.

Exercises:

8.

    a.

    b. Follow the steps outlined in <u>Detecting multicollinearity with simple correlation coefficients</u> and <u>Calculating Variance Inflation Factors</u> to check for high correlations and high VIF's in the implied regression model.

13. Follow the steps outlined in <u>Detecting multicollinearity with simple correlation coefficients</u> and <u>Calculating Variance Inflation Factors</u> to check *UE*, Equation 8.25, p. 269 for high correlations and high VIF's.

16.

    f. Run the regressions for this problem using the *Mine8.wf1* data set.

        a) Refer to <u>Chapter 5</u> for help with this part.

        b) Refer to <u>Detecting multicollinearity with simple correlation coefficients</u> for help with this part.

        c) Refer to <u>Transforming the multicollinear variables</u> for help with this part.

        d) Refer to <u>Table with EViews specification for functional forms</u> for help with this part.

Chapter 9: Serial Correlation

In this chapter:

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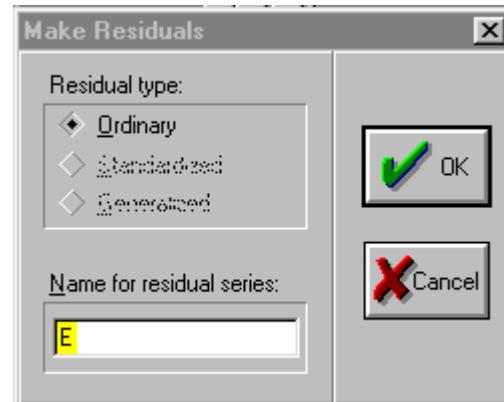
Plotting of the error term to detect serial correlation (*UE*, pp. 313-315):

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*.
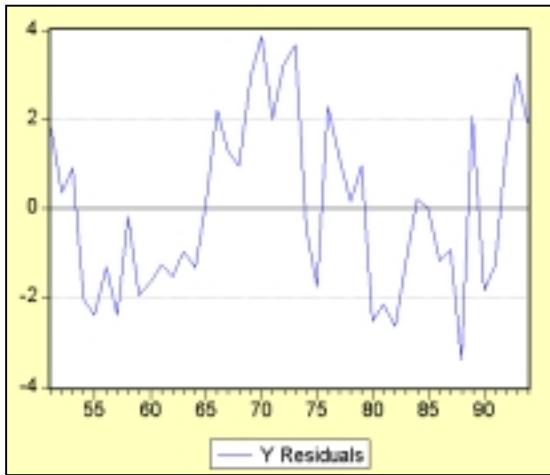
Step 3. Open the residual series named *E* in a new window by double clicking the series icon ☑ e in the workfile window to open the residual series from *EQ01* in a new window.

Step 4. Select **View/Line Graph** to reveal a time series graph of the residuals shown below.
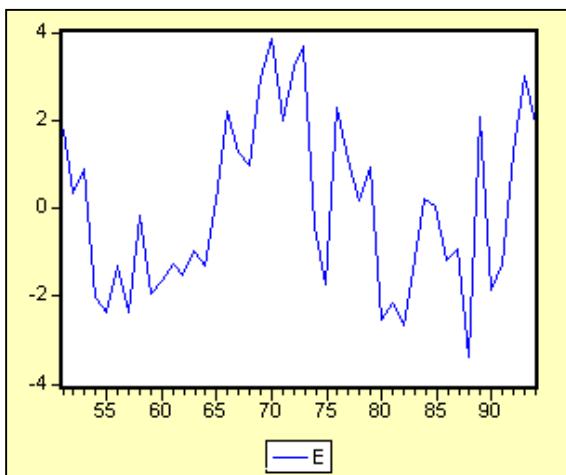
Complete the section entitled <u>Creating a residual series from a regression model</u> before attempting this section (i.e., Equation *EQ01* and series *E* should already be present in the workfile). Follow the steps below to estimate the first order serial correlation coefficient and test for possible 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 <u>N</u>ame 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).

Complete the section entitled <u>Creating a residual series from a regression model</u> 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/<u>E</u>stimation 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.

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 <u>Creating a residual series from a regression model</u> 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. <u>Estimate ρ, and name it *EQ02*</u> .
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 <u>Name to identify object:</u> 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]
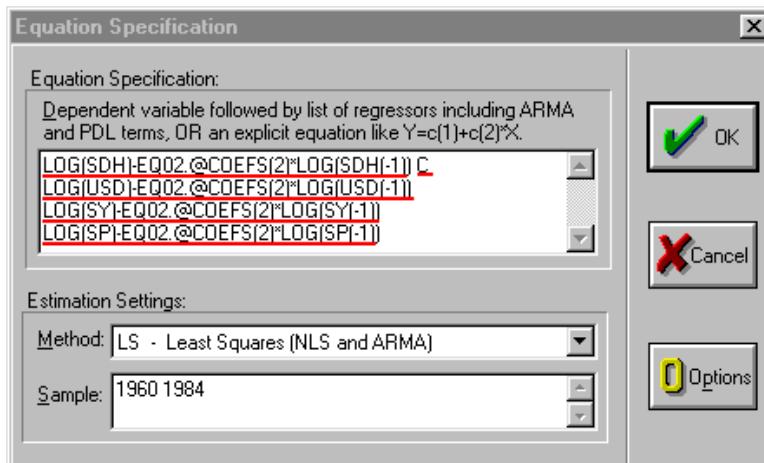
Step 5. To calculate the new residual series, enter the following formula in the command window: *series E = LOG(SDH)-(EQ03.@COEFS(1) + EQ03.@COEFS(2)*LOG(USD)*

---

[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 10th and 11th 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.

Chapter 10: Heteroskedasticity

The petroleum consumption example specified in *UE* 10.5 will be used to demonstrate all aspects of heteroskedasticity covered in this guide. Data for this problem is found in EViews workfile named *Gas10.wf1* and printed in Table 10.1 (*UE*, p. 371).

Graphing to detect heteroskedasticity (*UE* 10.1):

Figures 10.1, 10.2, and 10.3 in *UE* 10.1 demonstrate the value of a graph in detecting and identifying the source of heteroskedastic error. By graphing the residual from a regression against suspected variables, the researcher can often observe whether the variance of the error term changes systematically as a function of that variable.

Follow these steps to graph the residual from a regression against each of the independent variables in a model:

Step 1. Open the EViews workfile named *Gas10.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *PCON C REG TAX* in the Equation Specification: window, and. click **OK**.
Step 3. Select **Name** on the equation menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.
Step 4. Make a residual series named *E* and save the workfile.
Step 5. Make a Simple Scatter graph of *E against REG* to get the graph on the left below.
Step 6. Make a Simple Scatter graph of *E* against *TAX* to get the graph on the right below.
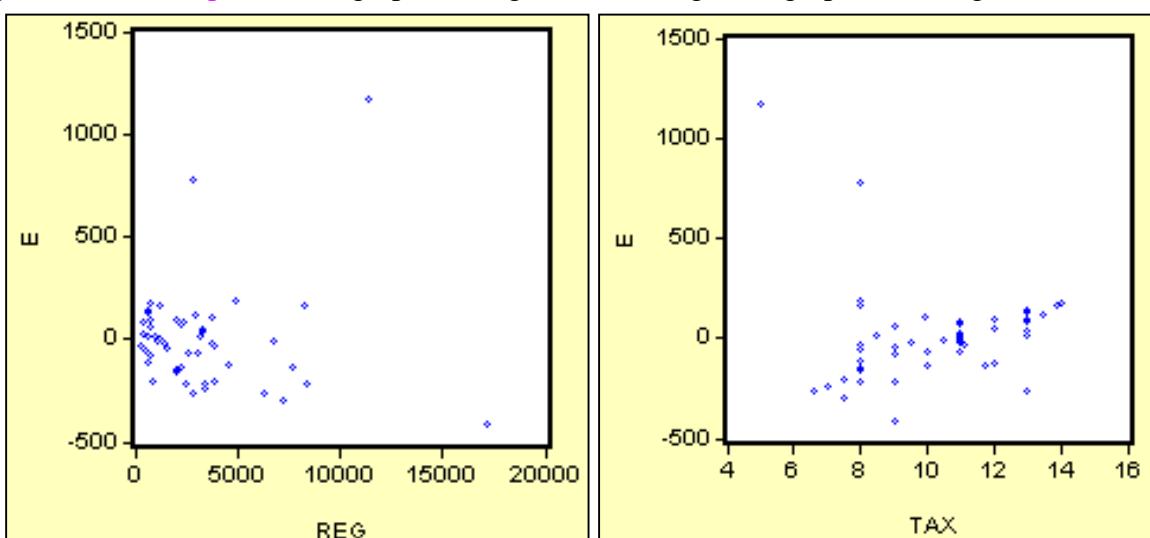
Testing for heteroskedasticity: the Park test (*UE* 10.3.2):

Complete **Steps 1-3** of the section entitled Graphing to detect heteroskedasticity before attempting this section. Follow these steps to complete the Park test for heteroskedasticity:

Step 1. Open the EViews workfile named *Gas10.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *log(E^2) C log(REG)* in the Equation Specification: window, and click **OK** to reveal the EViews output on the right.

Step 3. Test the significance of the coefficient on log(REG).

```
Dependent Variable: LOG(E^2)
Method: Least Squares
Date: 06/19/00   Time: 15:34
Sample: 1 50
Included observations: 50
```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.650264	2.374467	0.695004	0.4904
LOG(REG)	0.951920	0.308303	3.087609	0.0033

R-squared	0.165701	Mean dependent var	8.925457
Adjusted R-squared	0.148320	S.D. dependent var	2.248987
S.E. of regression	2.075512	Akaike info criterion	4.337470
Sum squared resid	206.7719	Schwarz criterion	4.413951
Log likelihood	-106.4368	F-statistic	9.533326
Durbin-Watson stat	1.759936	Prob(F-statistic)	0.003349

Testing for heteroskedasticity: White's test (*UE* 10.3.3 & *UE*, Equation 10.12):

Complete **Steps 1-3** of the section entitled Graphing to detect heteroskedasticity before attempting this section. Follow these steps to complete White's test for heteroskedasticity:

Step 1. Open the EViews workfile named *Gas10.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar and enter PCON C REG TAX in the Equation Specification: window. Click **OK**.

Step 3. To carry out White's heteroskedasticity test for the regression in **Step 2**, select **View/Residual Tests/White Heteroskedasticity (cross terms)** to get the output on the right. EViews reports two test statistics from the test regression. The Obs*R-squared statistic highlighted in yellow and boxed in red, is White's test statistic. It is computed as the number of

```
White Heteroskedasticity Test:
```

F-statistic	17.43051	Probability	0.000000
Obs*R-squared	33.22564	Probability	0.000003

```
Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 06/19/00   Time: 15:49
Sample: 1 50
Included observations: 50
```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1098291.	556502.1	1.973561	0.0547
REG	140.5628	39.92468	3.520699	0.0010
REG^2	-0.000515	0.001060	-0.485573	0.6297
REG*TAX	-12.84508	3.604180	-3.563939	0.0009
TAX	-237873.0	103340.2	-2.301844	0.0261
TAX^2	12346.99	4688.527	2.633449	0.0116

R-squared	0.664513	Mean dependent var	60168.93
Adjusted R-squared	0.626389	S.D. dependent var	207477.4
S.E. of regression	126817.9	Akaike info criterion	26.45106
Sum squared resid	7.08E+11	Schwarz criterion	26.68050
Log likelihood	-655.2764	F-statistic	17.43051
Durbin-Watson stat	2.250069	Prob(F-statistic)	0.000000

observations (n) times the R^2 from the test regression. White's test statistic is asymptotically distributed as a χ^2 with degrees of freedom equal to the number of slope coefficients, excluding the constant, in the test regression (five in this example).

Step 4. The critical χ^2 value can be calculated in EViews by typing the following formula in the EViews command window: $=@qchisq(.95,5)$.[1] After typing the formula and hitting **Enter** on the keyboard, `Scalar = 11.0704976935` appears in the lower left of the EViews screen (the same value found in *UE*, Table B-8). Since the nR^2 value of *33.2256393731* is greater than the 5% critical χ^2 value of *11.0704976935*, we can reject the null hypothesis of no heteroskedasticity. The probability printed to the right of the nR^2 value in the EViews output for White's heteroskedasticity test (i.e., *0.000003*) represents the probability that you would be incorrect if you rejected the null hypothesis of no heteroskedasticity.[2] The F-statistic is an omitted variable test for the joint significance of all cross products, excluding the constant. It is printed above White's test statistic for comparison purposes.

Remedies for heteroskedasticity: weighted least squares (*UE* 10.4.1):

Follow these steps to estimate the weighted least squares using *REG* as the proportionality factor:

Step 1. Open the EViews workfile named *Gas10.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *PCON/REG 1/REG REG/REG TAX/REG* in the Equation Specification: window, and click **OK**. Note the coefficients highlighted in yellow.

Dependent Variable: PCON/REG
Method: Least Squares
Date: 06/20/00   Time: 07:16
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1/REG	218.5395	48.10330	4.543129	0.0000
REG/REG	0.167831	0.013675	12.27294	0.0000
TAX/REG	-17.38900	4.682179	-3.713870	0.0005

R-squared	0.360003	Mean dependent var	0.198373
Adjusted R-squared	0.332770	S.D. dependent var	0.082386
S.E. of regression	0.067296	Akaike info criterion	-2.501302
Sum squared resid	0.212852	Schwarz criterion	-2.386581
Log likelihood	65.53256	F-statistic	13.21895
Durbin-Watson stat	1.921759	Prob(F-statistic)	0.000028

---

[1] $c=@qchisq(p,v)$ calculates the percentile of the χ^2 distribution . The formula finds the value c such that the prob(χ^2 with v degrees of freedom is $\leq c$) $= p$. In this case, the prob(χ^2 with 5 degrees of freedom is $\leq$ *11.0704976935*) $=95\%$. In other words, 95% of the area of a χ^2 distribution, with v=5 degrees of freedom, is in the range from *0* to *11.0704976935* and 5% is in the range from *11.0704976935* to ∞ (in the tail).

[2] The probability value is calculated with the formula $=@chisq(x,v)$, which returns the probability that a chi-squared statistic with v degrees of freedom exceeds x. To verify this, type the formula $=@chisq(33.2256393731,5)$ in the EViews command window and read the value of *3.39442964858e-06* on the lower left of the EViews screen.

Step 3. Select **Objects/New Object/Equation** on the workfile menu bar and enter *PCON C REG TAX* in the Equation Specification: window, and select the **O**Options button (see the arrow pointing toward the red box area in the graphic on the right).

Step 4. Check the **Weighted LS/TSLS** box and enter *1/REG* in the Weight: window (see the yellow highlighted and red boxed areas in the graphic on the right).

Step 5. Select **OK** to accept the options and select **OK** again to estimate the equation**.** Note that the weighted least squares coefficients found in **Step 2** are the same as the coefficients found in **Step 5** using the EViews weighted least squares option.[3]

Equation Specification

Equation Specification:
Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like Y=c(1)+c(2)*X.

PCON C REG TAX

OK
Cancel
Options

Estimation Settings:
Method: LS - Least Squares (NLS and ARMA)
Sample: 1 50

Estimation Options

LS and TSLS Options:
☐ Heteroskedasticity Consistent Covariance
◉ White
◇ Newey-West

☑ Weighted LS/TSLS (unavailable with ARMA)
Weight: 1/REG

Iterative procedures:
Max Iterations: 100
Convergence: 0.001

ARMA options:
Starting coefficient values
OLS/TSLS
☑ Backcast MA terms

OK
Cancel

---

[3] EViews performs weighted least squares by first dividing the weight series by its mean, then multiplying all of the data for each observation by the scaled weight series. The scaling of the weight series is a normalization that has no effect on the parameter results, but makes the weighted residuals more comparable to the un-weighted residuals. The normalization does imply, however, that EViews weighted least squares is not appropriate in situations where the scale of the weight series is relevant, as in frequency weighting.

Remedies for heteroskedasticity: heteroskedasticity corrected standard errors (*UE* 10.4.2):

Follow these steps to estimate heteroskedasticity corrected standard errors regression:

Step 1. Open the EViews workfile named *Gas10.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar and enter *PCON C REG TAX* in the Equation Specification: window, and select the ⬭Options button.

Step 3. Check the **Heteroskedasticity Consistent Covariances (White)** box (see the yellow highlighted and red boxed areas in the graphic on the right).

Step 4. Select **OK** to accept the options and select **OK** again to estimate the equation**.**

Step 5. Compare the Estimation Output from the regression with the Heteroskedasticity Consistent Covariance on the lower left with the Estimation Output from the uncorrected OLS regression on the lower right (*EQ01*). Note that the coefficients are the same but the uncorrected std. error is smaller. This means that the Heteroskedasticity Consistent Covariance correction has reduced the size of the t-statistics for the coefficients, a typical result. However, in this case both of the slope coefficients remain significant at the 5% level but the *TAX* variable coefficient is no longer significant at the 1% level.
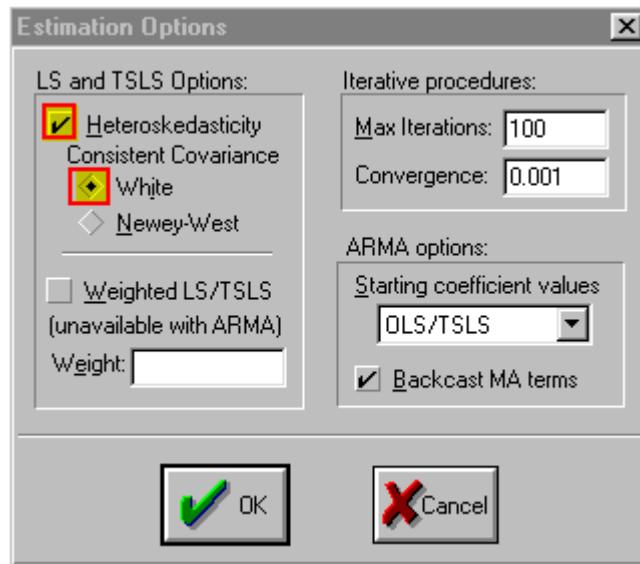
Dependent Variable: PCON
Method: Least Squares
Date: 06/20/00   Time: 08:28
Sample: 1 50
Included observations: 50
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	551.6880	237.1574	2.326252	0.0244
REG	0.186132	0.021533	8.644038	0.0000
TAX	-53.59101	23.89750	-2.242536	0.0297

R-squared	0.866368	Mean dependent var	603.7000
Adjusted R-squared	0.860682	S.D. dependent var	677.8267
S.E. of regression	253.0010	Akaike info criterion	13.96279
Sum squared resid	3008447.	Schwarz criterion	14.07751
Log likelihood	-346.0697	F-statistic	152.3567
Durbin-Watson stat	2.197170	Prob(F-statistic)	0.000000

Dependent Variable: PCON
Method: Least Squares
Date: 06/20/00   Time: 08:31
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	551.6880	186.2709	2.961750	0.0048
REG	0.186132	0.011719	15.88302	0.0000
TAX	-53.59101	16.85588	-3.179365	0.0026

R-squared	0.866368	Mean dependent var	603.7000
Adjusted R-squared	0.860682	S.D. dependent var	677.8267
S.E. of regression	253.0010	Akaike info criterion	13.96279
Sum squared resid	3008447.	Schwarz criterion	14.07751
Log likelihood	-346.0697	F-statistic	152.3567
Durbin-Watson stat	2.197170	Prob(F-statistic)	0.000000

Remedies for heteroskedasticity: redefining variables (*UE* 10.4.3):

Follow these steps to estimate *UE*, Equation 10.30, p. 374:

Step 1. Open the EViews workfile named *Gas10.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *PCON/POP C REG/POP TAX* in the Equation Specification: window, and click **OK**.

Exercises:

5. Create an EViews workfile and enter the average income and average consumption data from the table printed in Exercise 5, p.378.
    a. Refer to Testing for heteroskedasticity: the Park test.
    b. Refer to Testing for heteroskedasticity: the Park test.
    c.
    d. Refer to Remedies for heteroskedasticity: heteroskedasticity corrected standard errors.

9. Open the EViews file named *Books10.wf1*.
    a.
    b.
    c. Refer to Testing for heteroskedasticity: the Park test and Testing for heteroskedasticity: White's test.
    d. Refer to Remedies for heteroskedasticity: weighted least squares.
    e. Refer to Remedies for heteroskedasticity: heteroskedasticity corrected standard errors or Remedies for heteroskedasticity: redefining variables.

15. Open the EViews file named *Bid10.wf1*.
    a. Refer to Estimate a multiple regression model using EViews and Serial Correlation (Chapter 9).
    b. Refer to Serial Correlation (Chapter 9).
    c.
    d. Refer to Testing for heteroskedasticity: the Park test and Testing for heteroskedasticity: White's test.
    e. Refer to Testing for heteroskedasticity: the Park test and Testing for heteroskedasticity: White's test.
    f.

Chapter 11: A Regression User's Handbook

In this chapter:

How to observe checkpoint items displayed in *UE*, Table 11.1	
Checkpoint	**How to observe in EViews**
X, Y	Displaying the spreadsheet view of a group of variables
Degrees of Freedom (df)	To calculate the degrees of freedom for *EQ01*: enter the following formula in the command window, and press **Enter**: =*EQ01.@REGOBS-EQ01.@NCOEF* and view the df on the status line in the lower left of the screen
Coefficient (β-hat$_k$)	EViews Estimation Output table
t-Statistic (t)	EViews Estimation Output table
R-squared (R^2)	EViews Estimation Output table
Adjusted R-squared (R-bar^2)	EViews Estimation Output table
F-statistic (F)	EViews Estimation Output table
Durbin-Watson stat (DW)	EViews Estimation Output table
Residual (e_i)	Displaying the actual, fitted, residual, and a plot of the residuals, Graphing to detect heteroskedasticity
S.E. of regression (SEE)	EViews Estimation Output table
Total sum of squares (TSS)	To calculate the total sum of squares for a variable named *Y*: enter the following formula in the command window, and press **Enter**: series *SQUAREDERROR=(Y-@MEAN(Y))^2*. After that enter the following formula in the command window, and press **Enter**: =@SUM(*SQUAREDERROR*). The TSS can be viewed on the status line in the lower left of the screen.
Sum squared resid (RSS)	EViews Estimation Output table
Std. Error (of coefficients) {SE(β-hat$_k$)}	EViews Estimation Output table
Estimated first-order autocorrelation coefficient (ρ-hat))	Using regression to estimate ρ, the first order serial correlation coefficient
Simple correlation coefficient (r_{12})	Displaying the simple correlation coefficients between all pairs of variables in a group
Variance inflation factor (VIF)	Calculating Variance Inflation Factors

How to check for econometric maladies identified in *UE*, Table 11.2	
What's wrong	**How to check in EViews**
Omitted variable	Adding or deleting variables to/from an OLS model in EViews
Irrelevant variable	Adding or deleting variables to/from an OLS model in EViews
Incorrect functional form	Chapter 7 Specification: Choosing A Functional Form
Multicollinearity	Chapter 8: Multicollinearity
Serial correlation	Chapter 9: Serial Correlation
Heteroskedasticity	Chapter 10: Heteroskedasticity

Exercise:

UE 11.7.2, pp. 406-407:

Step 1. Open the EViews workfile named *House11.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *YOUR EQUATION SPECIFICATION HERE WITH THE DEPENDENT VARIABLE FIRST FOLLWOED BY C AND A LIST OF THE INDEPENDENT VARIABLE CHOSEN FOR EACH SPECIFICATION* in the Equation Specification: window, and. click **OK**.

Step 3. Check your results for each specification, following the outline printed in *UE*, p. 407 (refer to the tables above for guidance on how to perform each step).

Chapter 12: Time Series Models

In this chapter:
1. Estimating ad hoc distributed lag & Koyck distributed lag models (*UE* 12.1.3)
2. Testing for serial correlation in Koyck distributed lag models (*UE* 12.2.2) using:
   2.1. Durbin's h test
   2.2. The Lagrangian Multiplier (LM) test
3. Performing Granger Causality tests (*UE* 12.3.2)
4. Testing for nonstationarity by calculating the auto correlation function ACF (*UE* 12.4.1, Equation 12.24, p. 425)
5. Testing for nonstationarity with the Dickey-Fuller test (12.4.2)
6. Adjusting for nonstationarity (12.4.3)
7. Exercises

The workfile named *macro14.wf1* will be used to demonstrate the procedures explained in *UE*, Chapter 12. The examples examine the relationship between current purchases of goods and services (*CO*) and the level of disposable income (*YD*).

Estimating an ad hoc distributed lag model (*UE* 12.1.3):

To estimate the ad hoc distributed lag model printed in *UE*, Equation 12.14, follow these steps:

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *CO C YD(0 to -3)* in the Equation Specification: window, and click **OK**.[1]
Step 3. Select **Name** on the equation menu bar, enter the name *EQ01*, and click **OK**.
Step 4. Select **Save** on the workfile menu bar to save your changes.

Estimating a Koyck distributed lag model (*UE* 12.1.3):

To estimate the Koyck distributed lag model printed in *UE*, Equation 12.11, follow these steps:

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *CO C YD CO(-1)* in the Equation Specification: window, and click **OK.**
Step 3. Select **Name** on the equation menu bar, enter the name *EQ02*, and click **OK**.
Step 4. Select **Save** on the workfile menu bar to save your changes.

---

[1] You can include a consecutive range of lagged series by using the word "to" between the lags. *YD(0 to -3)* is equivalent to *YD YD(-1) YD(-2) YD(-3)*.

Testing for serial correlation in Koyck distributed lag models using Durbin's h test (*UE* 12.2.2):

Estimate the Koyck distributed lag model before attempting this section (i.e., Equation *EQ02* should already be present in the workfile). To conduct a Durbin's h test for *UE*, Equation 12.11, follow these steps:

Step 1. Open the EViews workfile named *Macro14.wf1*.

Step 2. To determine whether the value in parenthesis, in the denominator under the square root sign in *UE*, Equation 12.17, is positive, enter the following command in the command window:
   *scalar denominator=1-eq02.@regobs*(eq02.@stderrs(3)^2).*
Press **Enter** to create a scalar object named *denominator*. Double click the scalar object icon named *denominator* in the EViews workfile and view its value in the left corner of the status bar (bottom of the EViews window). If the number is positive, continue with the next step; if not, Durbin's h test is not valid.

Step 3. To compute Durbin's h test statistic shown in *UE*, Equation 12.17, enter the following command in the command window and press **Enter**:
       scalar dhtest=(1-(0.5*eq02.@dw))*sqr(eq02.@regobs/denominator).

Step 4. To view this scalar, double click the scalar object icon named *dhtest* and view its value in the left corner of the status bar (bottom of the EViews window). If the number is ≥ 1.96, reject the null hypothesis of no first order serial correlation.

Testing for serial correlation in Koyck distributed lag models using the Lagrangian Multiplier (LM) (*UE* 12.2.2):

Estimate the Koyck distributed lag model before attempting this section (i.e., equation *EQ02* should already be present in the workfile). To conduct a Lagrangian Multiplier (LM) test for *UE*, Equation 12.11, follow these steps:

Step 1. Open the EViews workfile named *Macro14.wf1*.

Step 2. Open the Equation named *EQ02* by double clicking its icon in the workfile window.

Step 3. Select **View/Residuals Tests/Serial Correlation LM Test…** on the equation menu bar (see highlighted selections in the graphic on the right).
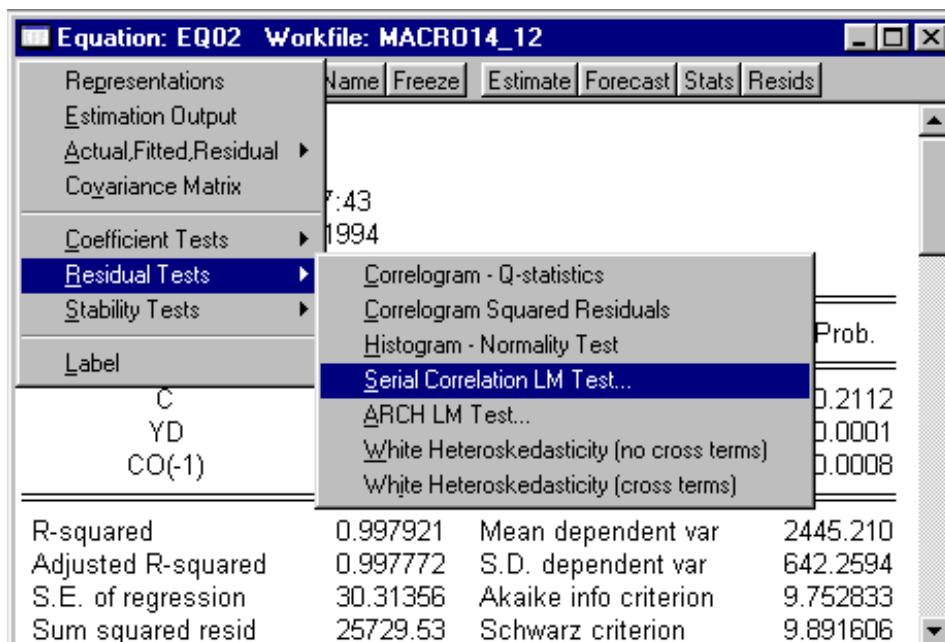
Step 4. Change the number in the Lags to include: to *1* in the Lag Specification: window.[2] Click **OK** to reveal the following EViews output:

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	11.77690	Probability		0.001946
Obs*R-squared	9.414982	Probability		0.002152
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Date: 07/01/00 Time: 10:25				
Presample missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-22.30297	26.12643	-0.853656	0.4008
YD	0.113797	0.104157	1.092547	0.2842
CO(-1)	-0.119235	0.110340	-1.080614	0.2894
RESID(-1)	0.602362	0.175526	3.431748	0.0019
R-squared	0.303709	Mean dependent var		3.67E-14
Adjusted R-squared	0.226343	S.D. dependent var		29.28568
S.E. of regression	25.75902	Akaike info criterion		9.455361
Sum squared resid	17915.24	Schwarz criterion		9.640392
Log likelihood	-142.5581	F-statistic		3.925632
Durbin-Watson stat	1.630947	Prob(F-statistic)		0.019027

The null hypothesis of the LM test is that there is no serial correlation up to lag order p, where p is equal to 1 in this case. The Obs*R-squared statistic is the Breusch-Godfrey LM test statistic. This LM statistic is computed as the number of observations times the R^2 from the test regression. The LM test statistic is asymptotically distributed as a χ^2 with p degrees of freedom (p is equal to 1 in this case).

Step 5. To determine whether the null hypothesis can be rejected in this case, determine the critical $\chi^2(1)$ value from *UE* Table B-8. The critical χ^2 value can also be calculated in EViews by typing the following formula in the EViews command window: $=@qchisq(.95,1)$.[3] EViews returns a scalar value of 3.84. Since the calculated Breusch-Godfrey LM test statistic of 9.42 exceeds the critical $\chi^2(1)$ value, we can reject the hypothesis of no serial correlation up to lag order 1 at the 95% confidence level. The probability printed to the right of the Obs*R-squared statistic in the EViews output (i.e., 0.002152) represents the probability that you would be incorrect if you rejected the null hypothesis of no serial correlation up to lag order 1 at the 95% confidence level.

---

[2] EViews enters 2 lags by default (i.e., testing for second order serial correlation). We will enter 1 lag to estimate the LM statistic for *UE*, Equation 12.20, p. 421.

[3] $c=@qchisq(p,v)$ calculates the percentile of the χ^2 distribution . The formula finds the value c such that the prob(χ^2 with v degrees of freedom is $\leq c) = p$. In this case, the prob(χ^2 with 1 degree of freedom is $\leq 3.84) = 95\%$. In other words, *95%* of the area of a χ^2 distribution, with *v=1* degrees of freedom, is in the range from *0* to *3.84* and *5%* is in the range from *3.84* to ∞ (i.e., in the tail of the distribution).

Performing Granger Causality tests (*UE* 12.3):

To conduct Granger Causality test for *CO* and *YD*, follow these steps:

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. To Create an EViews group for the current purchases of goods and services (*CO*) and the level of disposable income (*YD*), hold down the **Ctrl** button, click on, *CO* and *YD*, select **Show** from the workfile toolbar, and click **OK**.
Step 3. Select **Name** on the Group Object menu bar, enter *GROUP01* in the Name to identify object: window, and click **OK**.
Step 4. Select **View/Granger Causality…** on the Group Object menu bar. When you select the Granger Causality view, you will first see a dialog box asking for the number of lags to use in the test regressions. Change the number in the Lags to include: to *3* in the Lag Specification: window, and click **OK**.[4] EViews returns the following pairwise Granger Causality Tests Table.

Pairwise Granger Causality Tests			
Date: 07/01/00 Time: 14:43			
Sample: 1963 1994			
Lags: 3			
Null Hypothesis:	Obs	F-Statistic	Probability
YD does not Granger Cause CO	28	2.05100	0.13748
CO does not Granger Cause YD		4.57141	0.01291

Step 5. Based on the Probability values reported in the table, the hypothesis that *YD* does not Granger Cause *CO* cannot be rejected, but the hypothesis that *CO* does not Granger cause *YD* can be rejected. Therefore, it appears that Granger causality runs one way, from *CO* to *YD*, but not the other way.[5]

Testing for nonstationarity by calculating the auto correlation function ACF (*UE* 12.4.1, Equation 12.24, p. 425):

Follow these steps to calculate the auto correlation function ACF:

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. Open *CO* in one window by double clicking the series icon in the workfile window.
Step 3. To view the Autocorrelation and Partial Correlation, Select **View/Correlogram…**, on the *CO* series menu bar and a Correlogram Specification dialog box appears. Select *level* in the Correlogram of: window and enter *16* (the EViews default in this case) in the Lag Specification: lags to include: window, and click **OK** to reveal the EViews output below.

---

[4] In general, it is better to use more rather than fewer lags, since the theory is couched in terms of the relevance of all past information. You should pick a lag length that corresponds to reasonable beliefs about the longest time over which one of the variables could help predict the other.

[5] The reported F-statistics are the Wald statistics for the joint hypothesis that the coefficients on the lagged values of the other variable are zero for each equation (*CO* in the first equation and *YD* in the second equation for the table printed above). In case you want to determine significance by comparing the calculated F statistic with the critical F value from the F Table, the numerator degrees of freedom are given by the number of coefficient restrictions in the null hypothesis (i.e., the number of lags) and the denominator degrees of freedom are given by the total regression degrees of freedom.

Step 4. Since the AC's are significantly positive and the AC(k) dies off geometrically with increasing lag k, it is a sign that the series obeys a low-order autoregressive (AR) process.[6] In addition, since the partial autocorrelation (PAC) is significantly positive at lag 1 and close to zero thereafter, the pattern of autocorrelation can be captured by an autoregression of order one (i.e., AR(1)).

The finding in **Steps 3 & 4** indicates that the *CO*


```
Date: 07/08/00   Time: 09:36
Sample: 1963 1994
Included observations: 32
```

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.900	0.900	28.426	0.000
		2	0.804	-0.032	51.858	0.000
		3	0.713	-0.024	70.944	0.000
		4	0.628	-0.026	86.250	0.000
		5	0.535	-0.088	97.787	0.000
		6	0.444	-0.053	106.03	0.000
		7	0.353	-0.060	111.47	0.000
		8	0.266	-0.050	114.68	0.000
		9	0.182	-0.052	116.25	0.000
		10	0.106	-0.024	116.81	0.000
		11	0.040	-0.017	116.89	0.000
		12	-0.024	-0.052	116.92	0.000
		13	-0.082	-0.029	117.30	0.000
		14	-0.134	-0.040	118.38	0.000
		15	-0.182	-0.046	120.50	0.000
		16	-0.228	-0.059	124.04	0.000

series violates the third criteria for stationarity (*UE*, top of p. 425) and provides strong evidence that the *CO* series is non-stationary.

Testing for nonstationarity with the Dickey-Fuller (DF) test (12.4.2):

Since the AC analysis in the previous section indicated that *CO* is most likely an AR(1) process, the Dickey-Fuller (DF) test is valid. Follow these steps to conduct the Dickey-Fuller test of the hypothesis that the *CO* series is non-stationary:

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. Open *CO* in one window by double clicking the series icon in the workfile window. Note that EViews will probably display the correlogram view for *CO* since that was the last view selected in the previous section. You can select **View/Spreadsheet** to view the series data or just proceed with the next step to test for stationarity.
Step 3. To conduct the Dickey-Fuller (DF) test, select **View/Unit Root Test…** on the *CO* series window menu bar.
Step 4. Four things have to be specified in the Unit Root Test dialog box to carry out a unit root test. First, choose the type of test—either the Augmented Dickey-Fuller (ADF) test or the Phillips-

---

[6] If the AC(1) is nonzero, it means that the series is first order serially correlated. If AC(k) dies off more or less geometrically with increasing lag k, it is a sign that the series obeys a low-order autoregressive (AR) process. If AC(k) drops to zero after a small number of lags, it is a sign that the series obeys a low-order moving-average (MA) process. EViews also reports the Partial Correlations (PAC) in the same window. The partial correlation at lag k measures the correlation of *CO* values that are k periods apart, after removing the correlation from the intervening lags. If the pattern of autocorrelation is one that can be captured by an autoregression of order less than k, then the partial autocorrelation at lag k will be close to zero. The PAC of a pure autoregressive process of order k cuts off at lag k, while the PAC of a pure moving average (MA) process asymptotes gradually to zero.

Perron (PP) test (select *ADF* for this example).[7] Second, specify whether to test for a unit root in the Level, 1st difference, or 2nd difference of the series (select *level* for this example).[8] Third, specify whether to include an Intercept, a Trend and intercept, or None in the test regression. Select *Trend and intercept* for this example. To see why, read footnote 18, *UE*, p. 427. Fourth, specify the number of lagged first difference terms to add in the test regression (*0* for the DF test). The theory behind each of these selections is beyond the scope of *UE* and this guide. Advanced econometrics courses deal with these issues. When finished with the selections click **OK** to reveal the following table:

ADF Test Statistic	-1.633006	1% Critical Value*		-4.2826
		5% Critical Value		-3.5614
		10% Critical Value		-3.2138
*MacKinnon critical values for rejection of hypothesis of a unit root.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(CO)				
Method: Least Squares				
Date: 07/01/00 Time: 17:20				
Sample(adjusted): 1964 1994				
Included observations: 31 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
CO(-1)	-0.205313	0.125727	-1.633006	0.1137
C	320.1090	158.7047	2.017010	0.0534
@TREND(1963)	14.95367	8.798879	1.699497	0.1003
R-squared	0.107652	Mean dependent var		72.14839
Adjusted R-squared	0.043913	S.D. dependent var		38.54897
S.E. of regression	37.69307	Akaike info criterion		10.18860
Sum squared resid	39781.49	Schwarz criterion		10.32737
Log likelihood	-154.9232	F-statistic		1.688951
Durbin-Watson stat	1.191606	Prob(F-statistic)		0.202992

Step 5. The test fails to reject the null hypothesis of a unit root in the *CO* series at any of the reported significance levels, since the ADF Test Statistic[9] is not less than (i.e., does not lie to the left of) the MacKinnon critical values.

---

[7] EViews refers to both the Dickey-Fuller and the Augmented Dickey-Fuller tests as ADF tests. You will face two practical issues in performing the ADF test. First, you will have to specify the number of lagged first difference terms to add to the test regression (selecting zero yields the DF test; choosing numbers greater than zero generates ADF tests). The usual (though not particularly useful) advice is to include lags sufficient to remove any serial correlation in the residuals. Second, EViews asks you whether to include other exogenous variables in the test regression. You have the choice of including a constant, a constant and a linear time trend, or neither in the test regression.

[8] You can use this option to determine the number of unit roots in the series. If the test fails to reject the test in levels but rejects the test in first differences, then the series contains one unit root and is of integrated order one I(1). If the test fails to reject the test in levels and first differences but rejects the test in second differences, then the series contains two unit roots and is of integrated order two I(2).

[9] The output reports the ADF Test Statistic, but in reality, it is the DF test statistic, since zero lags were chosen.

Adjusting for nonstationarity (12.4.3):

In order to determine whether the first differenced series[10] is stationary, follow the steps in the previous section and select *1st difference* for the Test for unit root in: window and *Intercept* in the Include in test equation: window. Note that the null hypothesis of a unit root in the first difference of *CO* can be rejected at the 5% but not at the 1% level. This adds to the evidence from the ACF test that indicates *CO* is most likely an AR(1) process.

Exercises:

5. Open the EViews workfile named *Mouse12.wf1*.
   a. Follow the steps in estimating distributed lag models.
   b. Follow the steps in estimating Koyck lag models.

6. Open the EViews workfile named *Mouse12.wf1*.
   a. Complete Exercise 5b and follow the steps found in Testing for serial correlation in Koyck lag models using Durbin's h test.
   b. Complete Exercise 5b and follow the steps found in Testing for serial correlation in Koyck lag models using the Lagrangian Multiplier (LM) test.

7. Open the EViews workfile named *Mouse12.wf1*. Complete Exercise 5b and follow the steps found in using the Lagrangian Multiplier (LM) test to detect serial correlation tests in Koyck lag models.
   a.
   b.
   c. In **Step 5**, change the number in the Lags to include: to *2* in the Lag Specification: window.

8. Open the EViews workfile named *Macro14.wf1* and follow the steps in the Performing Granger Causality tests section for *I* and *Y* (instead of *CO* and *YD*).

10. Open the EViews workfile named *Macro14.wf1* and follow the steps in the Testing for nonstationarity by calculating the auto correlation function ACF section.

11. Open the EViews workfile named *Macro14.wf1* and follow the steps in the Testing for nonstationarity with the Dickey-Fuller test section.

---

[10] To use first differencing to rid a series of nonstationarity, simply enter *D(CO)* for the series in any EViews procedure. For example, entering *D(CO)* as the dependent variable in a least squares regression is the same as entering *CO-CO(-1)*.

Chapter 13: Dummy Dependent Variable Techniques

In this chapter:
1. Estimating the linear probability model (*UE* 13.1.3)
2. Estimating the Weighted Least Squares (WLS) correction for heteroskedasticity in the linear probability model (*UE* 13.1.3, pp. 441-442)
3. Estimating the binomial logit model (*UE* 13.2)
4. Estimating the binomial probit model (*UE* 13.3.1)
5. Interpreting the results of binary dependent variable regression
6. Estimating the multinomial logit model (*UE* 13.3.2)
7. Exercises

The workfile named *women13.wf1* will be used to demonstrate the procedures explained in *UE*, Chapter 13. The data for this example are printed in *UE*, Table 13.1, p. 440. The name of the dummy variable is changed from *D*, in *UE*, Table 13.1, to *j* in the *Women13.wf1* workfile (D is a reserved name in EViews).

Estimating the linear probability model (*UE* 13.1.3):

To estimate the linear probability model printed in *UE*, Equation 13.6, follow these steps:

Step 1. Open the EViews workfile named *Women13.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *J C M S* in the Equation Specification: window, and click **OK**.

Step 3. Select **Name** on the equation menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.

Step 4. Select **Forecast** on the equation menu bar, enter *JFOLS* in the Forecast name: window, and click **OK**.

Step 5. Enter the formula *series JFOLSP = JFOLS >= 0.5* in the command window and press **Enter**. A series named *JFOLSP* is created that predicts whether a women is expected to be in the labor force based on the linear probability model. The formula applies the decision rule: *JFOLSP* is *1* if, the probability value, $JFOLS_i \geq 0.5$ and *JFOLSP* is *0* if $JFOLS_i < 0.5$.

Step 6. Enter the formula *series OLSP = JFOLSP=J* in the command window and press **Enter** to calculate a series that equals 1 if the OLS model predicted correctly and 0 if not.

Step 7. Enter the formula *scalar R2pOLS = @sum(OLSP)/@obs(OLSP)* in the command window and press **Enter** to calculate the R^2_p value for the OLS model. Double click the ⊞ r2pols icon to reveal the percentage of correct predictions from the OLS model on the status line in the lower left of the screen (0.8 for this exercise).

Step 8. Select **Save** on the workfile menu bar to save your changes.

Estimating the Weighted Least Squares (WLS) correction for heteroskedasticity in the linear probability model (*UE 13.1.3, pp. 441-442*):

To estimate the weighted least squares model specified in *UE*, Equations 13.7 & 13.8, follow these steps:

Step 1. Open the EViews workfile named *women13.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *J/Z C 1/Z M/Z S/Z* in the Equation Specification: window, and click **OK** to generate the same coefficients and standard errors reported in *UE*, Equation 13.8.

Step 3. Select **Name** on the equation menu bar, enter *EQ02a* in the Name to identify object: window, and click **OK**.

Step 4. To perform the same analysis using the WLS feature in EViews: select **Objects/New Object/Equation** on the workfile menu bar, enter *J Z C M S* in the Equation Specification: window, select the ⬭Options button, check Weighted LS/TSLS and type *1/Z* in the Weight: window. Click **OK** twice to generate the same coefficients and standard errors reported in *UE*, Equation 13.8. Note that the coefficient on the Z variable is the constant (i.e., $Z*(1/Z) = 1$) and the coefficient on the constant, *C*, is the coefficient on the variable 1/Z (i.e., $C*(1/Z)$).

Step 5. Select **Name** on the equation menu bar, enter *EQ02b* in the Name to identify object: window, and click **OK**.

Step 6. Select **Forecast** on the equation menu bar, enter *JFWLS* in the Forecast name: window, and click **OK**.

Step 7. Enter the formula *series JFWLSP = JFWLS >= 0.5* in the command window and press **Enter**. A series named *JFWLSP* is created that predicts whether a women is expected to be in the labor force based on the linear probability model. The formula applies the decision rule: *JFWLSP* is *1* if, the probability value, $JFWLS_i \geq 0.5$ and *JFWLSP* is *0* if $JFWLS_i < 0.5$.

Step 8. Enter the formula *series WLSP =JFWLSP=J* in the command window and press **Enter** to calculate a series that equals 1 if the WLS model predicted correctly and 0 if not.

Step 9. Enter the formula *scalar R2pWLS = @sum(WLSP)/@obs(WLSP)* in the command window and press **Enter** to calculate the R^2_p value for the WLS model. Double click the *R2pWLS* icon to reveal the percentage of correct predictions from the WLS model (0.83 for this exercise).

Step 10. Select **Save** on the workfile menu bar to save your changes.

Estimating the binomial logit model (*UE 13.2*):

To estimate the binomial logit model printed in *UE*, Equation 13.15, follow these steps:

Step 1. Open the EViews workfile named *women13.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar.

Step 3. From the Estimation Settings:, Method: window, select the *BINARY - Binary choice (logit, probit, extreme value...)* estimation method. The window will change to reflect your choice.

Step 4. There are two parts to the binary model specification. First, in the Equation Specification: field, you should type the name of the Binary dependent variable followed by a list of regressors (i.e., enter *J C M S* in the Equation Specification: window for this example). Second, check *logit* as the Binary estimation method: (this is the default setting in EViews). Click **OK** to run the logit regression.

Step 5. Select **Name** on the equation menu bar, enter *EQ03* in the Name to identify object: window, and click **OK**.

Step 6. Select **Forecast** on the equation menu bar, enter *JFLOG* in the Forecast name: window, and click **OK**.

Step 7. Enter the formula *series JFLOGP = JFLOG >= 0.5* in the command window and press **Enter**. A series named *JFLOGP* is created that predicts whether a women is expected to be in the labor force based on the linear probability model. The formula applies the decision rule: *JFLOGP* is *1* if, the probability value, $JFLOG_i \geq 0.5$ and *JFLOGP* is *0* if $JFLOG_i < 0.5$.

Step 8. Enter the formula *series LOGP = JFLOGP=J* in the command window and press **Enter** to calculate a series that equals 1 if the LOG model predicted correctly and 0 if not.

Step 9. Enter the formula *scalar R2pLOG = @sum(LOGP)/@obs(LOGP)* in the command window and press **Enter** to calculate the R^2_p value for the LOG model. Double click the *R2pLOG* icon to reveal the percentage of correct predictions from the LOG model (0.8 for this exercise).

Step 10. Select **Save** on the workfile menu bar to save your changes.

The linear probability model results and the binomial logit model results can be compared by opening both regression equation results in the work area (i.e., double click the *EQ01* and *EQ03* equation icons in the workfile window.

Estimating the binomial probit model (*UE* 13.3.1):

To estimate the binomial probit model printed in *UE*, Equation 13.19, follow these steps:

Step 1. Open the EViews workfile named *women13.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar.

Step 3. From the Equation Specification: window, select the *BINARY - Binary choice (logit, probit, extreme value…)* estimation method. The window will change to reflect your choice.

Step 4. There are two parts to the binary model specification. First, in the Equation Specification: field, you should type the name of the Binary dependent variable followed by a list of regressors (i.e., enter *J C M S* in the Equation Specification: window for this example). Second, check *probit* as the Binary estimation method: (*logit* is the default setting in EViews). Click **OK** to run the probit regression.

Step 5. Select **Name** on the equation menu bar, enter*EQ04* in the Name to identify object: window, and click **OK**.

Step 6. Select **Forecast** on the equation menu bar, enter *JFPRO* in the Forecast name: window, and click **OK**.

Step 7. Enter the formula *series JFPROP = JFPRO >= 0.5* in the command window and press **Enter**. A series named *JFPROP* is created that predicts whether a women is expected to be in the labor force based on the linear probability model. The formula applies the decision rule: *JFPROP* is *1* if, the probability value, $JFPRO_i \geq 0.5$ and *JFPROP* is *0* if $JFPRO_i < 0.5$.

Step 8. Enter the formula *series PROP = JFPROP=J* to calculate a series that equals 1 if the PRO model predicted correctly and 0 if not.

Step 9. Enter the formula *scalar R2pPRO = @sum(PROP)/@obs(PROP)* to calculate the R^2_p value for the PRO model. Double click the *R2pPRO* icon to reveal the percentage of correct predictions from the PRO model (0.8 for this exercise).

Step 10. Select **Save** on the workfile menu bar to save your changes.

**Interpreting the results of binary dependent variable regression:**

The estimated coefficient on each independent variable is easy to interpret in an OLS model, but difficult to interpret in a model estimated using the probit or logit technique. However, the relative size of each coefficient reflects the relative effect of the independent variables on the predicted probability for the dependent variable. Interpretation of the coefficient values is complicated by the fact that estimated coefficients from a binary dependent model cannot be interpreted as the marginal effect on the dependent variable.

Estimating the multinomial logit model (*UE* 13.3.2):

The multinomial logit process will not be discussed in detail here because the EViews 3.1 Student Version does not have the capability of estimating such programs.

Exercises:

8.
    a.
    b. EViews gives you the error message Near singular matrix when the logit is used to estimate the model with the Equation Specification: *J C M S AD* (see graphic on the right.

9. Follow the procedures outlined in:
    a. Estimating the linear probability model.
    b. Estimating the binomial logit model.
    c. Estimating the binomial probit model.

12. Open the EViews workfile named *Mort13.wf1*.
    a.
    b. See Estimating the linear probability model.
    c. See Estimating the linear probability model.
    d.
    e. See Estimating the linear probability model and Estimating the binomial logit model.
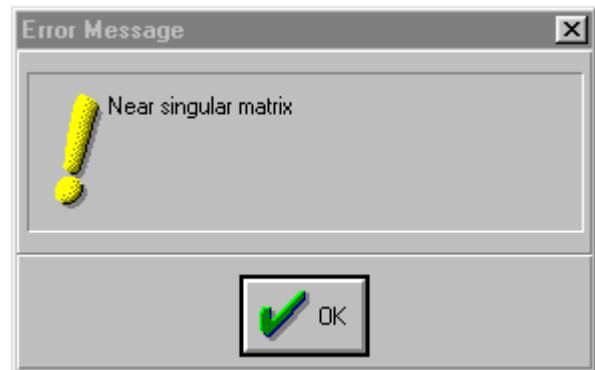
Chapter 14: Simultaneous Equations

The naïve Keynesian macroeconomic model of the U.S. economy identified in *UE*, p. 477 will be used to demonstrate the two stage-least squares procedure. The data for this model is found in the EViews workfile named *macro14.wf1* and it is printed in *UE*, Table 14.1, p. 478. Two variables that are included in the macroeconomic model must be generated from other data series (see note at the bottom of *UE*, Table 14.1, p. 478).

Generating time series for taxes and net exports using structural equations (*UE*, p. 477):

Follow these steps to generate time series values for T (taxes) and NX (net exports) using the structural equations in the model:

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. To generate a new series named *T* for taxes, select **Genr** on the workfile menu bar, type *T=Y-YD* in the Enter equation: window, and click **OK**. A new series icon for *T* is created in the workfile window.
Step 3. To generate a new series named *NX* for net exports, select **Genr** on the workfile menu bar, type *NX=Y-CO-I-G* in the Enter equation: window, and click **OK**. A new series icon for *NX* is created in the workfile window.
Step 4. Select **Save** on the workfile menu bar to save your changes.

Estimating *CO* with least squares (*UE*, Equation 14.31, p. 481):

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *CO C YD CO(-1)* in the Equation Specification: window, and click **OK** to reveal the regression output to the right.
Step 3. Select **Name** on the equation window menu bar, enter *OLS_CO* in the Name to identify object: window, and click **OK**.
Step 4. Select **Save** on the workfile menu bar to save your changes.

Dependent Variable: CO
Method: Least Squares
Date: 07/11/00   Time: 07:50
Sample(adjusted): 1964 1994
Included observations: 31 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-38.10541	29.77949	-1.279586	0.2112
YD	0.516486	0.116196	4.444959	0.0001
CO(-1)	0.461118	0.123244	3.741502	0.0008

|			
R-squared	0.997921	Mean dependent var	2445.210
Adjusted R-squared	0.997772	S.D. dependent var	642.2594
S.E. of regression	30.31356	Akaike info criterion	9.752833
Sum squared resid	25729.53	Schwarz criterion	9.891606
Log likelihood	-148.1689	F-statistic	6719.462
Durbin-Watson stat	0.892667	Prob(F-statistic)	0.000000

Estimating two-stage least squares regression using EViews TSLS method (*UE*, 14.3.1):

To estimate the two-stage least squares model printed in *UE*, Equation 14.29, follow these steps:

Step 1. Open the EViews workfile named *Macro14.wf1*.

Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, and select *TSLS – Two-Stage Least Squares (TSNLS and ARMA)* in the Method: window under Estimation Settings: and the dialog will change to include an Instrument list: window (see graphic on the right).

Step 3. Enter CO C YD CO(-1) in the Equation Specification: window and *C G T NX CO(-1) R(-1)* in the Instrument list: window.[1] The graphic above shows the relevant selections/entries highlighted in yellow. Click **OK** to reveal the Estimation Output view printed below. The yellow highlighted portions of the regression output reflect the selections made in the dialog window shown above.[2]
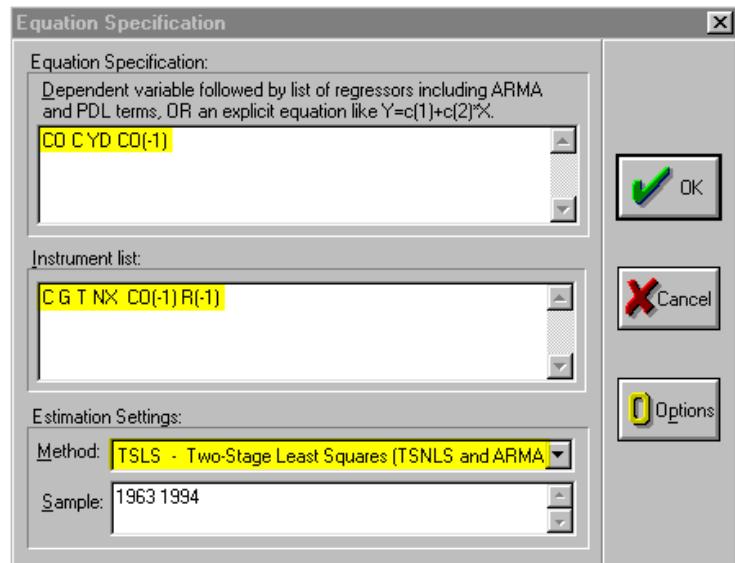
Dependent Variable: CO				
Method: Two-Stage Least Squares				
Date: 07/10/00 Time: 15:12				
Sample(adjusted): 1964 1994				
Included observations: 31 after adjusting endpoints				
Instrument list: C G T NX CO(-1) R(-1)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-24.73014	34.90233	-0.708553	0.4845
YD	0.441638	0.153839	2.870773	0.0077
CO(-1)	0.540309	0.163000	3.314782	0.0025
R-squared	0.997890	Mean dependent var		2445.210
Adjusted R-squared	0.997739	S.D. dependent var		642.2594
S.E. of regression	30.53734	Sum squared resid		26110.82
F-statistic	6615.725	Durbin-Watson stat		0.982576
Prob(F-statistic)	0.000000			

Step 4. Select **Name** on the equation window menu bar, enter *TSLS_CO* in the Name to identify object: window, and click **OK**.

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

---

[1] The constant, C, is always a suitable instrument, so EViews will add it to the instrument list if you omit it.

[2] EViews identifies the estimation procedure, as well as the list of instruments in the header. This information is followed by the usual coefficient, t-statistics, and asymptotic p-values. EViews uses the structural residuals in calculating all of the summary statistics. These structural residuals should be distinguished from the second-stage residuals that you would obtain from the second-stage regression if you actually computed the two-stage least squares estimates in two separate stages.

Estimating two-stage least squares regression using two distinct stages and OLS (*UE*, 14.3.1):

   To estimate the two-stage least squares equation printed in *UE*, Equation 14.28, using ordinary OLS and two distinct phases, follow these steps:

Step 1. Open the EViews workfile named *Macro14.wf1*.

Step 2. To estimate the reduced form equation for *YD* (*UE*, Equation 14.27, p. 480), select **Objects/New Object/Equation** on the workfile menu bar, enter *YD C G NX T CO(-1) R(-1)* in the Equation Specification: window, and click **OK**.

Step 3. To generate the forecast values from this equation, select **Forecast** on the equation menu bar, enter *YDF* in the Forecast name: window, and click **OK**. EViews will create a new variable in the workfile named *YDF*.

Step 4. To estimate the second stage equation for *CO* (*UE*, Equation 14.29, p. 481), select **Objects/New Object/Equation** on the workfile menu bar, enter *CO C YDF CO(-1)* in the Equation Specification: window, and click **OK.** Note that we have used the instrumental variable *YDF* instead of the actual variable *YD* for disposable income. The method, dependent variable, and variable names are highlighted in yellow in the OLS regression output shown below.

Step 5. Select **Name** on the equation window menu bar, enter *TSLS_OLS_CO* in the Name to identify object: window, and. click **OK**.

Step 6. Select **Save** on the workfile menu bar to save your changes.

Dependent Variable: CO
Method: Least Squares
Date: 07/05/00   Time: 15:44
Sample(adjusted): 1964 1994
Included observations: 31 after adjusting endpoints

| Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-24.73014	41.09577	-0.601769	0.5522
YDF	0.441638	0.181138	2.438126	0.0214
CO(-1)	0.540309	0.191924	2.815219	0.0088
R-squared	0.997075	Mean dependent var		2445.210
Adjusted R-squared	0.996866	S.D. dependent var		642.2594
S.E. of regression	35.95622	Akaike info criterion		10.09425
Sum squared resid	36199.78	Schwarz criterion		10.23302
Log likelihood	-153.4608	F-statistic		4771.906
Durbin-Watson stat	1.485932	Prob(F-statistic)		0.000000

Comparing the OLS, EViews TSLS, and OLS two-stage models:

To compare the coefficients, std. Errors, and t-statistics for the three models discussed in this chapter, open the equations named *OLS_CO*, *TSLS_CO* and *TSLS_OLS_CO*, by double clicking their respective icons in the workfile window, and compare the regression output. To facilitate the process, the output for the OLS, EViews TSLS and OLS TSLS models are printed in this guide. Look at all three and compare the data printed in the red-boxed area for each regression. Note that the estimated coefficients are larger in the *OLS_CO* model compared to the *TSLS_CO* and *TSLS_OLS_CO* models. This supports the hypothesis that OLS estimates of coefficients have a positive bias in simultaneous equation models (simultaneity bias). Contrarily, TSLS estimated coefficients tend to have a downward bias. Note that the estimated coefficients are identical for the *TSLS_CO* and *TSLS_OLS_CO* models, but the standard errors (Std. Error in the EViews output) are smaller in the EViews TSLS estimated model, making the coefficients more significant (i.e., higher t-statistics). In order to get accurate estimates of standard errors and t-scores, the estimation should be done on a complete two-stage least squares program (like EViews TSLS). When OLS is used to estimate the second stage, it ignores the fact that the first stage was run at all (*UE*, footnote 11, p. 481).

The identification problem and the order condition (*UE*, 14.3.3):

In order to calculate two-stage least squares using the *TSLS – Two-Stage Least Squares (TSNLS and ARMA)* option, your specification must satisfy the order condition for identification, which states that there must be at least as many instruments as there are coefficients in your equation. The order condition for identification is easy to assess in EViews. Count, to make sure that the number of independent variables, not counting the constant, in the Equation Specification: window (i.e., *YD & CO(-1)*) is less than or equal to the number of predetermined variables in the Instrument list: window (i.e., *G, T, NX, CO(-1) & R(-1)*). See graphic in the Two-stage least squares regression using EViews TSLS method section above.

Exercises:

9. Open EViews and open the EViews workfile named *Macro14.wf1*.
   a. Refer to [Estimating *CO* with least squares](#).
   b. Refer to **Step 2** of [Estimating two-stage least squares regression using two distinct stages and OLS](#).
   c. Refer to **Steps 3** & **4** of [Estimating two-stage least squares regression using two distinct stages and OLS](#).
   d. Refer to [Estimating two-stage least squares regression using EViews TSLS method](#).

12. Double click the ▣ tsls_co icon in the EViews *Macro14.wf1* workfile window to re-activate the *UE*, Equation 14.29. Click **Estimate** on the equation menu bar and click **OK**. The reason for this is to make sure that the residuals in the EViews workfile are from the *TSLS_CO* equation. If the ▣ tsls_co icon is not in the workfile, you must go back and follow the steps outlined in [Estimating two-stage least squares regression using EViews TSLS method](#).
   a. Follow the procedures outlined in [Chapter 9](#).
   b.

13. Open EViews and open the EViews workfile named *Oats14.wf1*.
   a.
   b.
   c.
   d. Refer to [Estimating *CO* with Least Squares (OLS)](#) and [Estimating two-stage least squares regression using EViews TSLS method](#).
   e. Refer to [Comparing the OLS, EViews TSLS, and OLS two-stage models](#).

Chapter 15: Forecasting

Forecasting chicken consumption using OLS (*UE* 15.1, Equation 6.8, p. 501):

The chicken demand model developed in Chapter 6 was estimated using data from 1951-1994. In order to forecast a variable beyond 1994, the workfile range and sample must be expanded. Follow these steps to forecast chicken consumption for 1995 - 1997 using ordinary least squares:

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

Step 2. To expand the workfile range, select **Procs/Change Workfile Range** from the workfile menu bar, change the End date from 1994 to 1997, and click **OK**.

Step 3. To expand the sample, select **Sample** on the workfile menu bar, change the second number in the window from 1994 to 1997, and click **OK**. After **Steps 2 & 3** are completed, the top portion of the EViews workfile should look like the graphic on the right (note that the second date for Range: & Sample: have changed from 1994 to 1997).

Step 4. To open *Y, PC, PB, & YD* in a group window, hold down the **Ctrl** button, click on *Y, PC, PB, & YD*, select **Show** from the workfile toolbar, and click **OK**.

Step 5. To enter the data for *Y, PC, PB, & YD* from the text & footnote number 4 in *UE*, p. 502 (see table below), click **edit+/-** on the group menu bar, scroll to the bottom of the spreadsheet and replace the *NA*'s with the appropriate numbers for the period 1995 - 1997. Press **Enter** after each entry and click **edit+/-** on the group menu bar a second time to save your changes. Scroll to the bottom of the group spreadsheet and make sure that it looks like the table below.

	Y	*PC*	*PB*	*YD*
1995	80.30000	6.500000	61.80000	200.6200
1996	81.90000	6.700000	58.70000	208.5000
1997	83.70000	7.700000	63.10000	216.3100

Step 6. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C PC PB YD* in the Equation Specification: window, change the Sample: to 1951 - 1994, and click **OK**.[1]

Step 7. Select **Name** on the equation window menu bar, enter *EQ01* in the Name to identify object: window, and click **OK**.

Step 8. Select **Forecast** on the equation menu bar, enter *YFOLS* in the Forecast name: window, set the Sample range to forecast: to 1951 1997, and click **OK**.

Step 9. Open *YFOLS & Y* in a group window by holding down the **Ctrl** button, clicking on *YFOLS & Y*, selecting **Show** from the workfile toolbar, and clicking **OK**. Scroll to the bottom of the group spreadsheet and make sure that it looks like the table below. Note that the predicted values for *YFOLS* in the EViews spreadsheet are a slightly different than those in the *UE*, p. 502 table, because the text uses rounded coefficient values for Equation 6.8 and we used non-rounded EViews estimated coefficients.

	YFOLS	Y
1995	80.71830	80.30000
1996	82.06108	81.90000
1997	83.65985	83.70000

Step 10. Select **Save** on the workfile menu bar to save your changes.

Forecasting chicken consumption using a generalized least squares model estimated with the AR(1) method (*UE*, 15.2.2):

Complete the section entitled Forecasting chicken consumption using OLS before attempting this section. The OLS estimate for chicken consumption should already have been estimated and saved as *EQ01* in the *Chick6.wf1* workfile and the workfile range & sample should have been expanded to 1997 and the data for 1995 - 1997 entered into the workfile spreadsheet.

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, change the Sample: to 1951 - 1994, and click **OK**.

Step 3. Select **Name** on the equation window menu bar, enter *EQ04* in the Name to identify object: window, and click **OK**.

Step 4. Select **Forecast** on the equation menu bar, enter *YFAR1* in the Forecast name: window, set the Sample range to forecast: to 1951 1997, and click **OK**.

Step 5. Open *YFAR1 & Y* in a group window by holding down the **Ctrl** button, clicking on *YFOLS & Y*, selecting **Show** from the workfile toolbar, and clicking **OK**. Scroll to the bottom of the group spreadsheet and make sure that it looks like the table below.

	YFAR1	Y
1995	80.05606	80.30000
1996	81.67044	81.90000
1997	83.85621	83.70000

Step 6. Select **Save** on the workfile menu bar to save your changes.

---

[1] The reason for changing the workfile sample back to the original setting is to ensure that the equation estimation uses only the data in the original sample (i.e., 1951-1994).

Forecasting chicken consumption using a generalized least squares model estimated with the Cochrane-Orcutt method(*UE*, 15.2.2):

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 a GLS model for chicken consumption. If you have questions concerning the procedure, review the appropriate section of Chapter 9. Complete the section entitled Forecasting chicken consumption using OLS before attempting this section. The OLS estimate for chicken consumption should already have been estimated and saved as *EQ01* in the *Chick6.wf1* workfile and the workfile range & sample should have been expanded to 1997 and the data for 1995 - 1997 entered into the workfile spreadsheet.

Step 1. Open the EViews workfile named *Chick6.wf1*, select **Sample** on the workfile menu bar, change the End date from 1997 to 1994, and click **OK**.

Step 2. Open *EQ01* by double clicking its icon in the workfile window. Create a new series for the residuals (errors) for *EQ01*, by selecting **Procs/Make Residual Series** on the *EQ01* window menu bar, enter the name *E* as the Name for residual series, and click **OK**.

Step 3. Select **Objects/New Object/Equation** on the workfile menu bar. Enter *E C E(-1)* in the Equation Specification: window and click **OK**. The coefficient on the *E(-1)* term (i.e., ρ), is positive and significant at the 1% level (t-statistic = 3.69 and Prob value = 0.0006). This evidence points to positive serial correlation. Select **Name** on the equation menu bar, enter *EQ02* in the Name to identify object: window, and click **OK**.

Step 4. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y-EQ02.@COEFS(2)*Y(-1) C PC-EQ02.@COEFS(2)*PC(-1) PB-EQ02.@COEFS(2)*PB(-1) YD-EQ02.@COEFS(2)*YD(-1)* in the Equation Specification: window, and click **OK**. Select **Name** on the equation menu bar, enter *EQ03* in the Name to identify object: window, and click **OK**.[2]

Step 5. Calculate the new residual series by typing the following formula in the command window: *series E = Y-(EQ03.@COEFS(1) + EQ03.@COEFS(2)*PC + EQ03.@COEFS(3)*PB + EQ03.@COEFS(4)*YD)* and pressing **Enter** on the keyboard.

Step 6. Re-run *EQ02, EQ03* and the *series E* equation[3] in step 5 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 6 iterations, the value for ρ converged (i.e., ρ changed from 0.902484 to 0.902802 between the 5th and 6th 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 pressing **Enter** on the keyboard.

Step 8. Double click the ▦ beta0 icon in the workfile window and read the value, *26.72580*, for the estimated constant in the lower left of the screen. The final GLS equation for chicken consumption is *Y = 26.72580 - 0.109891*PC + 0.090293*PB + 0.242032*YD*, similar to the truncated coefficient model printed in *UE*, Equation 9.22, p. 507.

---

[2] Note that 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. The equation, with the entire variable names printed out, can be viewed by selecting **View/Representations** on the equation menu bar.

[3] 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.

Step 9. Select **Sample** on the workfile menu bar and change the End date from 1994 to 1997. Open *EQ03* and select **Forecast** on the equation menu bar. Make sure that *Y* is checked in the Forecast of: window. Change the Forecast name: to *YFGLS*. Check to make sure that the Sample range to forecast: is set to 1951 1997, and click **OK** (see graphic on the right).

Step 10. Open *YFGLS & Y* in a group window by holding down the **Ctrl** button, clicking on *YFGLS & Y*, selecting **Show** from the workfile toolbar, and clicking **OK**. Scroll to the bottom of the group spreadsheet and make sure that it looks like the table below.

	YFGLS	Y
1995	80.05495	80.30000
1996	81.66931	81.90000
1997	83.85514	83.70000

Step 11. Select **Save** on the workfile menu bar to save your changes.
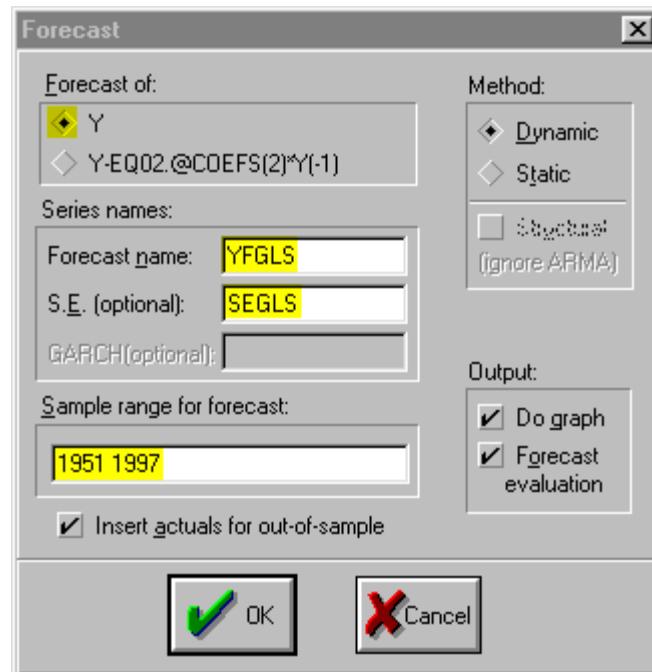
Step 12. Note that the predicted values for *YFGLS* in the EViews spreadsheet are slightly different than those in the *UE*, p. 502 table. This is due to the fact that the text uses rounded coefficient values for Equation 9.22 and we used non-rounded EViews estimated values. Delete this group object when finished.

Forecasting confidence intervals (*UE*, 15.2.3):

Follow these steps to make a point prediction of the weight of a male who stands 6'1" tall and calculate the high and low estimated values for a 95% confidence interval.

Step 1. Open the EViews workfile named *Htwt1.wf1*.

Step 2. Follow **Steps 2 & 3** of the section entitled Forecasting chicken consumption using OLS to expand the workfile range and sample range from *20* to *21*. Double click the *X* variable icon in the workfile window, click **edit+/-** on the series menu bar, scroll to the bottom of the spreadsheet, replace the *NA* for the 21st observation with 13 (i.e., height in inches above 5'), and press **Enter**. To save your changes, click **edit+/-** on the series menu bar a second time.

Step 3. Select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C X* in the Equation Specification: window, and click **OK**. Select **Name** on the equation menu bar and enter *EQ01* in the Name to identify object: window.

Step 4. Select **Forecast** on the equation menu bar. Enter *YF* in the Forecast name: window. Check to make sure that the Sample range to forecast: is set to 1 21, and click **OK**. To view the forecast weight of the male student standing 6'1" tall, double click the *YF* series icon in the workfile window and scroll to the bottom. The forecast weight for the 21st observation is 186.2993.

Step 5. Select **Sample** on the workfile menu bar, change the Sample range pairs (or sample object to copy): to 1 20, and click **OK**.

Step 6. Select **Procs/Make Residual Series** on the *EQ01* window menu bar. Enter the name *E* as the Name for residual series, and click **OK**.

Step 7. Generate a new series for the residuals squared (i.e., *E2*) by selecting **Genr** on the workfile menu bar, entering the equation: *E2=E^2* in the Enter equation: window, and clicking **OK.**

Step 8. Generate a new series named *XDEV2* for the residuals squared by selecting **Genr** on the workfile menu bar, entering the equation: *XDEV2=(x-@mean(x))^2* in the Enter equation: window, and clicking **OK**.

Step 9. To calculate the upper confidence interval for the 6'1" student, type the following formula (all one equation) in the command window, and press **Enter** on the keyboard:
scalar YF_HIGH=Yf(21)+((@sum(E2)/@obs(E2)(1+(1/@obs(E))+((X(21)-@mean(X))^2/@sum(XDEV2))))^.5)*@qtdist(.975,(eq01.@regobs-eq01.@ncoef))*

Double click *YF_HIGH* in the workfile window to view the upper confidence interval in the lower left of the screen (i.e., ☐ Scalar YF_HIGH = 204.279137482).
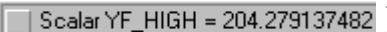
Step 10. To calculate the lower confidence interval for the 6'1"student, type the following formula (all one equation) in the command window and press **Enter** on the keyboard:
scalar YF_LOW=Yf(21)-((@sum(E2)/@obs(E2)(1+(1/@obs(E))+((X(21)-@mean(X))^2/@sum(XDEV2))))^.5)*@qtdist(.975,(eq01.@regobs-eq01.@ncoef))*

Double click *YF_LOW* in the workfile window to view the upper confidence interval in the lower left of the screen (i.e., ☐ Scalar YF_LOW = 168.319457872)
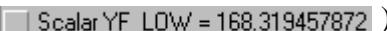
Forecasting with simultaneous equation systems (*UE* 15.2.4):

EViews has a Type of Object: named **Model** that enables you to solve a system of simultaneous equations and use the model for forecasting and simulation. EViews models do not contain unknown coefficients to be estimated. Instead, the **Model** object allows you to solve for unknown values for the endogenous variables. For a description of various methods of creating model objects in EViews and techniques for using these objects to forecast and perform simulations see **Help/Contents/Hypothesis Testing and Forecasting/Model Solve (Forecast and Simulation)**.

This topic is beyond the scope of this guide but, if you were to do it, you would follow these steps to forecast with a simultaneous equation model:

Step 1. Create a model by selecting **Object/New Object/Model** in the main toolbar, type in a name for your model in the Name for Object: window, click **OK**, and enter the previously estimated equations in the model window.

Step 2. To solve the model, simply select the **Solve** button in the model toolbar. You should see the Model Solution dialog box offering various options for controlling the solution process. EViews solves for the endogenous variables, given data for the exogenous variables.

Forecasting with ARIMA models (*UE* 15.3):

ARIMA (autoregressive, integrated, moving average) models use three tools for modeling the serial correlation in the disturbance:
1. The first tool is the autoregressive or AR term. The AR(1) model introduced above uses only the first-order term but, in general, you may use additional, higher-order AR terms. Each AR term corresponds to the use of a lagged value of the residual in the forecasting equation for the unconditional residual. An autoregressive model of order p is denoted as AR(p).
2. The second tool is the integration order term. Each integration order corresponds to differencing the series being forecast. A first-order integrated component means that the forecasting model is designed for the first difference of the original series. A second-order component corresponds to using second differences, and so on.
3. The third tool is the MA, or moving average term. A moving average forecasting model uses lagged values of the forecast error to improve the current forecast. A first-order moving average term uses the most recent forecast error, a second-order term uses the forecast error from the two most recent periods, and so on.

In ARIMA forecasting, you assemble a complete forecasting model by using combinations of the three building blocks described above. You can use the correlogram view of a series for this purpose (see Testing for nonstationarity by calculating the auto correlation function ACF for a description of this process).

UE did not present an example of ARIMA estimation, but one will be estimated here to demonstrate how to estimate an ARIMA model in EViews. In Chapter 12 we determined that *CO* was integrated order one and probably followed an AR(1) process. For purposes of illustration, follow these steps to estimate an ARIMA(*1,1,2*)[4] model for *CO*:

Step 1. Open the EViews workfile named *Macro14.wf1*.
Step 2. Select **Objects/New Object/Equation** on the workfile menu bar, enter *D(CO) C AR(1) MA(1) MA(2)* in the Equation Specification: window, and click **OK** to view the EViews Estimation Output printed below.

Step 3. Select **Name** on the equation menu bar and enter *EQ01* in the Name to identify object: window.

Step 4. Follow **Steps 2 & 3** of the section entitled Forecasting chicken consumption using OLS to expand the workfile range and sample range from *1994* to *1998*.

Step 5. Select **Forecast** on the *EQ01* menu bar. Enter *COF* in the Forecast name: window. Check to make sure that the Sample range to forecast: is set to 1998, and click **OK**.

Dependent Variable: D(CO)				
Method: Least Squares				
Date: 07/15/00 Time: 10:35				
Sample(adjusted): 1965 1994				
Included observations: 30 after adjusting endpoints				
Convergence achieved after 100 iterations				
Backcast: OFF (Roots of MA process too large for backcast)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	67.74008	3.655560	18.53070	0.0000
AR(1)	0.718868	0.251279	2.860836	0.0082
MA(1)	-0.943222	0.362834	-2.599598	0.0152
MA(2)	-0.773132	0.292764	-2.640801	0.0138
R-squared	0.588152	Mean dependent var		72.04333
Adjusted R-squared	0.540631	S.D. dependent var		39.20346
S.E. of regression	26.57083	Akaike info criterion		9.521071
Sum squared resid	18356.23	Schwarz criterion		9.707897
Log likelihood	-138.8161	F-statistic		12.37671
Durbin-Watson stat	2.201209	Prob(F-statistic)		0.000032
Inverted AR Roots	.72			
Inverted MA Roots	1.47	-.53		
Estimated MA process is noninvertible				

Step 6. To view the forecast, double click the *COF* series and scroll to the bottom of the spreadsheet to view the following forecast values for 1995 - 1998:

	COF
1995	3536.473
1996	3604.213
1997	3671.953
1998	3739.694

---

[4] ARIMA(*1,1,2*) represents a model that has a first order autoregressive term (i.e., AR(1)), is integrated order one and has two moving average terms (i.e., MA(1) & MA(2)).

2.
   a.
   b. Refer to <u>Running a simple regression for Woody's Restaurants example</u> to estimate Equation 3.7 and <u>Forecasting chicken consumption using OLS</u> to make an unconditional forecast of the gross sales volume at the three proposed restaurant sites.
   c. Open the EViews workfile named *Fish8.wf1*, select **Objects/New Object/Equation** on the workfile menu bar, enter *F C PF/PB log(YD) P* in the Equation Specification: window, and click **OK**. Then follow the steps in <u>Forecasting chicken consumption using OLS</u> to make an unconditional forecast of *F* for 1971-1974, given the numbers printed in the table in *UE*, Exercise 2c, p.517.

3. Open the EViews workfile named *Htwt1.wf1*, select **Objects/New Object/Equation** on the workfile menu bar, enter *Y C X* in the Equation Specification: window, and click **OK**. Ask three students their height (*X=the height in inches exceeding 5'*). Then follow the steps in <u>Forecasting chicken consumption using OLS</u> to make the Unconditional forecast of the weight (*YF*) for each student in your sample. Ask these students their actual weight (*Y*) and compare *Y* with *YF*.

4.
   a. Follow the steps in <u>Forecasting confidence intervals</u> to calculate the 95% confidence interval for the weight for a person standing 5'9" tall (i.e., *X = 9*).
   b.

6. Follow the steps in <u>Chapter 1</u> for each part of this exercise.

Chapter 16: Statistical Principles

In this chapter:

Describing data (*UE* 16.1):

The annual percentage returns for 25 stock mutual funds printed in *UE*, Table 16.2, p. 522 will be used to show the EViews commands needed to calculate the descriptive statistics described in *UE* 16.1.

Follow these steps to view a histogram and the standard descriptive statistics for a series:

Step 1. Create a new Undated or irregular workfile with 25 observations. Use **Genr** to create a new variable named *Y1997* and enter the 1997 returns from *UE*, Table 16.2, p. 522 into the series. Repeat the process for 1998 returns (name the variable *Y1998*). Refer to Chapter 1 for help.

Step 2. Save the workfile by selecting **File/Save As…** on the EViews main menu bar, and enter *Mutual16.wf1* in the File name: window (check to make sure *workfile* is selected in the Save as type: window).

Step 3. Open the variable named *Y1997* in a new window by double clicking its icon in the workfile window.

Step 4. Select **View/Descriptive Statistics/Histogram and Stats** on the workfile menu bar to reveal the graphic shown below (statistics highlighted in yellow are described in *UE* 16.1).
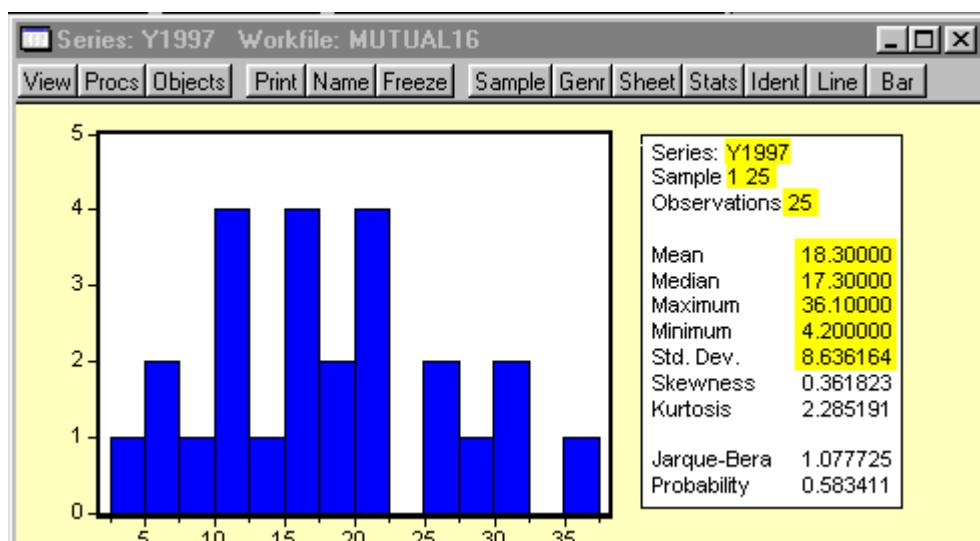
The EViews output shows a histogram of the data series plus major descriptive statistics. The histogram divides the series range (the distance between the maximum and minimum values) into a number of equal length intervals or bins and displays a count of the number of observations that fall into each bin. The histogram is useful when investigating the distributional characteristics of a series (see *UE* 16.2).

The descriptive statistics window on the right of the graphic displays the standard descriptive statistics. All of the statistics are calculated using observations in the current sample. The top of the window identifies the ==series name==, ==sample==, and the ==number of observations==. The descriptive statistics described in *UE* 16.1 are presented next and they are defined as follows:

1. ==Mean== (the average value of the series, obtained by adding up the series and dividing by the number of observations)
2. ==Median== (the middle value (or average of the two middle values) of the series when the values are ordered from the smallest to the largest)
3. ==Maximum== (the maximum value of the series in the current sample)
4. ==Minimum== (the minimum value of the series in the current sample)
5. ==Std. Dev.== (standard deviation) is a measure of dispersion or spread in the series

Individual scalar values for the following descriptive statistics[1] can be obtained by typing the functions, printed in the first column of the table below, in the command window, and pressing **Enter**. The values can be viewed on the status line in the lower left portion of the screen.

Function	Name
=@obs(Y1997)	number of observations
=@mean(Y1997)	mean
=@median(Y1997)	median
=@min(Y1997)	minimum
=@max(Y1997)	maximum
=@stdev(Y1997)	sample standard deviation
=@sum(Y1997)	sum
=@sumsq(Y1997)	sum-of-squares
=@var(Y1997)	variance
=@var(Y1997)*(@obs(Y1997)/(@obs(Y1997)-1))	sample variance =@stdev(Y1997)^2

Note that EViews calculates the population variance, which is a biased measure of the population variance in samples. Since the variance is calculated as the sum of the squared deviations of a series observations about its mean divided by the number of observations (i.e., divided by *n* instead of an unbiased estimator, which divides by *n-1*), it is better to calculate the sample variance as the square of the unbiased sample standard deviation {i.e., $=@stdev(Y1997)\^2$}. Due to the way EViews calculates variance, think of the variance as the square of the standard deviation instead of the other way around.

---

[1] The variable *Y1997* is used for demonstration purposes so you can try it out in the *Mutual16.wf1* workfile. Substitute any variable name, for *Y1997*, to calculate the statistic for that variable in a workfile. Other measures obtainable by EViews commands can be found in **Help/Function Reference** under the heading: Descriptive Statistics Functions.

The remaining four numbers displayed in the descriptive statistics window (see the graphic on the first page) are defined below:

1. Skewness (the skewness of a symmetric distribution, such as the normal distribution, is zero)
2. Kurtosis (the kurtosis of the normal distribution is 3)
3. Jarque-Bera test (under the null hypothesis of a normal distribution, the Jarque-Bera statistic is distributed as χ^2 with 2 degrees of freedom
4. Probability (the probability that a Jarque-Bera statistic exceeds the observed value under the null indicates that a small probability value leads to the rejection of the null hypothesis of a normal distribution)

These measures of normality are not discussed in *UE* but the normal distribution and its properties are described in *UE* 16.2.5, pp. 539 - 543.

The same descriptive statistics can be calculated for a group of variables, sans the histogram, by opening a group of variables in one window and selecting **View/Descriptive Statistics/Histogram and Stats** on the workfile menu bar.

Probability distributions (*UE* 16.2):

EViews enables you to calculate the cumulative distribution density (CDF or inverse CDF) or probability functions, cumulative distribution, and random number generators for 17 statistical distributions. We have already used EViews to calculate the critical t-value for t-tests (see Chapter 5 and Calculating a confidence interval for a population mean) and generated random numbers for the Monte Carlo Simulation to demonstrate that the estimated βs are drawn from a normal distribution (see Chapter 4). Further discussion of this topic is beyond the scope of this guide, but further a explanation of EViews capabilities relating to statistical distribution functions can be found in **Help/Function Reference**.

Standardized variables (*UE* 16.2.4):

Complete **Steps 1 & 2** of Describing data prior to attempting this section. To calculate the standardized values for *Y1997*, follow these steps:

Step 1. Open the EViews workfile named *Mutual16.wf1*.
Step 2. Type *series Y1997standized = (y1997-@mean(Y1997))/@stdev(Y1997)* in the command window and press **Enter** on the keyboard. Y1997standized successfully computed. will appear in the status line in the lower left of the screen.
Step 3. To view the standardized values for *Y1997*, double click the *Y1997standized* series icon in the workfile menu.

Calculating a confidence interval for a population mean (*UE* 16.4.6):

Complete **Steps 1 & 2** of Describing data before attempting this section. To calculate the 95%[2] confidence interval for the population mean of *Y1998*:

---

[2] To compute the 99% confidence interval, substitute *.995* for *.975* in **Step 2** and **Step 3** below.

Step 1. Open the EViews workfile named *Mutual16.wf1*.

Step 2. To calculate the upper confidence interval, type *scalar CI_Y1998_HIGH = @mean(Y1998)+(@qtdist(.975,@obs(Y1998)-1)*(@stdev(Y1998)/(@obs(Y1998)^.5)))* in the command window, and press **Enter**. Double click on the *CI_Y1998_HIGH* icon in the workfile to view the value on the status line in the lower left portion of the screen. Did you get ☐ Scalar CI_Y1998_HIGH = 14.1972846046 ?

Step 3. To calculate the lower confidence interval, type *scalar CI_Y1998_LOW = @mean(Y1998)-(@qtdist(.975,@obs(Y1998)-1)*(@stdev(Y1998)/(@obs(Y1998)^.5)))* in the command window, and press **Enter**. Double click on the *CI_Y1998_LOW* icon in the workfile to view the value on the status line in the lower left portion of the screen. Did you get ☐ Scalar CI_Y1998_LOW = 3.65871539542 ?

Hypothesis testing, the test statistic, and statistical significance (*UE* 16.5.3 & *UE* 16.5.4):

Complete steps 1 and 2 of Describing data prior to attempting this section. To test the hypothesis that the average return for all mutual funds in 1998 was 28.1 percent based on the sample of 25 mutual fund returns (i.e., series *Y1998*), follow these steps:

Step 1. Open the EViews workfile named *Mutual16.wf1*.

Step 2. Double click the *Y1998* series icon in the workfile window.

Step 3. Select **View/Tests for Descriptive Stats/Simple Hypothesis Tests** on the series menu bar to reveal the Series Distributions Tests dialog window. Enter *28.*1 in the Mean: window under Test Value: and click **OK** to reveal the EViews test output table shown below.[3]

Hypothesis Testing for Y1998			
Date: 07/16/00 Time: 16:29			
Sample: 1 25			
Included observations: 25			
Test of Hypothesis: Mean = 28.10000			
Sample Mean = 8.928000			
Sample Std. Dev. = 12.76537			
Method		Value	Probability
t-statistic		-7.509381	0.0000

EViews prints values for the: Test of Hypothesis: Mean, Sample mean, Sample Std. Dev., t-statistic: Value (based on the formula for t printed in *UE*, p. 558), and Probability. The t-statistic value of -7.509381 is equal to the value printed in *UE*, p. 558. This value can be compared with the critical t-value found in the *UE*, Table B-1 using the one-tailed or two-tailed test at various levels of confidence. The reported Probability is the p-value, or marginal significance level, against a two-sided alternative. If this probability value is less than the size of the test, say 0.05, we reject the null hypothesis (i.e., Mean = 28.10000). The probability value for a one-sided alternative is one-half the p-value of the two-sided test.

---

[3] You can enter a value for the series standard deviation in the window under Mean Test Assumption:, if it is known. If unknown, leave the window blank and EViews will use the sample standard deviation in the test calculation.

Exercises:

3. Follow the steps in [Describing data](#) to calculate the mean and standard deviation.

10. Follow the steps in [Describing data](#) to create a workfile for the problem. Then follow the steps in [Calculating a confidence interval for a population mean](#) to calculate the 99% confidence interval (check footnote 2). You could also answer this problem by following the procedures in [Hypothesis testing, the test statistic, and statistical significance](#).

13. Follow the steps in [Describing data](#) to create a workfile for the problem. Then follow the steps in [Hypothesis testing, the test statistic, and statistical significance](#) to calculate the t-value and probability that the sample is drawn from a population with a mean value of 11.2.