In this chapter, you will focus on:

- The concept of autocorrelation: Errors in a regression model should not follow a pattern.
- Different types of autocorrelation.
- Symptoms of autocorrelation.
- How to test for autocorrelation.
- Designing models so that there is no autocorrelation in the first place.
- How to force autocorrelation out if necessary.

Suppose that every morning you watch a weather report featuring your favorite meteorologist on the local morning news. You use his forecasts to decide what clothes to wear. You notice that in the winter months, once you get outside, you are always cold. After a while, you realize that the temperature predicted by the meteorologist is always higher than the actual temperature. Then, during the spring, you realize the temperature he predicts is always lower than the actual temperature. Something is wrong with the meteorologist’s method. Since his winter forecasts are always too high, why doesn’t he start lowering his predictions? When spring rolls around, then his predictions are always too low. Why doesn’t he realize this and raise his spring forecasts? This is the essence of autocorrelation: The errors follow a pattern, showing that something is wrong with the regression model.

One of the classical assumptions of the ordinary least squares procedure is that the observations of the error term are independent of each other. Each error term observation must not be correlated with the error term observation that is next to it. If this assumption is violated and the error term observations are correlated, autocorrelation is present. Autocorrelation is a common problem in time-series regressions. Like other violations of the classical assumptions, we can view autocorrelation as a regression “illness.” When autocorrelation is present, the error term observations follow a pattern. Such patterns tell us that something is wrong.

In Section 7-1, we examine why autocorrelation is a problem for ordinary least squares. In Section 7-2, you will learn how to recognize the symptoms of autocorrelation. Next
we present a test for autocorrelation. The final two sections of the chapter describe two possible solutions to the problem. Section 7-4 discusses how you can prevent the illness, by eliminating the underlying cause of the autocorrelation. Section 7-5 shows you how to treat the symptoms by eliminating the observed effects of autocorrelation, even if you cannot cure the illness itself.

7-1 THE ILLNESS

**Autocorrelation** (also called *serial correlation*) occurs when the error term observations in a regression are correlated. The theoretical error term $e$ is a random variable that is part of the regression model, even before it is estimated (see Section 1-2). This error term represents a random “shock” to the model, or something that is missing from the model. However, we can never see the actual error term $e$. Therefore, we use the error term observations or residuals ($\hat{e}$) to check for autocorrelation. If they follow a pattern, this pattern is evidence of autocorrelation.

Remember that the error term observations (residuals) are $Y - \hat{Y} = \hat{e}$. We see the error term observations only after we estimate the model, and they differ with each data set used to estimate the same regression. In this way, the difference between $e$ and $\hat{e}$ parallels the difference between $B$ and $\hat{B}$. (The true value of the slope $B$ can never be observed; we find the slope estimate $\hat{B}$ only after we estimate a regression. The value for the slope estimate $\hat{B}$ differs when different data are used to estimate the same regression.)

We will use a simple model examining the relationship between Microsoft’s marketing and advertising expenditures and its revenues to illustrate autocorrelation. The regression is:

\[
\text{REVENUES} = B_0 + B_1 \text{MARKETING} + B_2 \text{SUMMER} + B_3 \text{FALL} + B_4 \text{WINTER} + e
\]  

(7-1)

where

- **REVENUES** = Microsoft’s real quarterly revenues, in millions of dollars
- **MARKETING** = Microsoft’s real quarterly expenditures on marketing and advertising, in millions of dollars
- **SUMMER** = 1 when REVENUES and MARKETING are from the third quarter (July-September), 0 otherwise
- **FALL** = 1 when REVENUES and MARKETING are from the fourth quarter (Oct.-Dec.), 0 otherwise
- **WINTER** = 1 when REVENUES and MARKETING are from the first quarter (Jan.-March), 0 otherwise

Table 7-A gives the data from the first quarter of 1987 through the third quarter of 2000. We will present the regression results using this data in a later section. Now, however, let’s look at the error term observations from those regression results (Figure 7-1).

The top square plotted in Figure 7-1 is the observed error for the first quarter in 1987; the bottom square is the observed error for the third quarter of 2000. Error term

---

1 “Real” means the values are adjusted for inflation.
### Table 7-A
Data for Microsoft Revenue Model

<table>
<thead>
<tr>
<th>Year</th>
<th>Quarter</th>
<th>REVENUES</th>
<th>MARKETING</th>
<th>SUMMER</th>
<th>FALL</th>
<th>WINTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1987</td>
<td>1</td>
<td>60.02</td>
<td>13.44</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>71.62</td>
<td>16.80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>85.66</td>
<td>18.36</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>86.68</td>
<td>22.54</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1988</td>
<td>1</td>
<td>88.74</td>
<td>23.26</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>132.73</td>
<td>31.48</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>136.02</td>
<td>32.75</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>141.27</td>
<td>40.72</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1989</td>
<td>1</td>
<td>145.48</td>
<td>37.81</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>169.81</td>
<td>46.90</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>158.02</td>
<td>40.91</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>175.58</td>
<td>45.29</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1990</td>
<td>1</td>
<td>183.55</td>
<td>49.99</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>232.73</td>
<td>56.44</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>237.14</td>
<td>58.53</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>252.80</td>
<td>64.32</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1991</td>
<td>1</td>
<td>273.74</td>
<td>72.70</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>339.97</td>
<td>95.87</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>357.80</td>
<td>94.39</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>383.44</td>
<td>96.59</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1992</td>
<td>1</td>
<td>421.15</td>
<td>115.38</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>490.70</td>
<td>138.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>486.87</td>
<td>130.59</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>578.58</td>
<td>155.74</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

(continued)
### Table 7-A (Continued)

<table>
<thead>
<tr>
<th>Year</th>
<th>Quarter</th>
<th>REVENUES</th>
<th>MARKETING</th>
<th>SUMMER</th>
<th>FALL</th>
<th>WINTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>1</td>
<td>575.82</td>
<td>165.62</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>655.34</td>
<td>199.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>669.35</td>
<td>195.49</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>718.96</td>
<td>190.72</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1994</td>
<td>1</td>
<td>678.94</td>
<td>174.51</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>776.25</td>
<td>188.</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>846.69</td>
<td>201.43</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>876.14</td>
<td>200.49</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1995</td>
<td>1</td>
<td>841.80</td>
<td>212.11</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>996.06</td>
<td>265.44</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1064.33</td>
<td>285.22</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1082.03</td>
<td>263.02</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1996</td>
<td>1</td>
<td>1345.16</td>
<td>331.61</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1461.03</td>
<td>353.92</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1468.54</td>
<td>363.48</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1493.38</td>
<td>343.85</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1997</td>
<td>1</td>
<td>1507.21</td>
<td>324.63</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1752.81</td>
<td>398.88</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2092.23</td>
<td>397.31</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2079.69</td>
<td>380.88</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1998</td>
<td>1</td>
<td>2059.30</td>
<td>416.92</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2329.72</td>
<td>468.15</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2438.19</td>
<td>429.62</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2532.22</td>
<td>420.21</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
observations to the left of the center line are less than zero and the observed errors to the right of the line are greater than zero. Here a positive observed error is usually followed by another positive observed error. Likewise, a negative observed error is usually followed by a negative observed error. The sign of the preceding observed error thus predicts the sign of almost all of the observed errors shown in Figure 7-1. These observed errors follow an easily recognized pattern. They are not random. This signals that our model has an autocorrelation problem.

The presence of autocorrelation does not mean that the values of one independent variable are correlated over time. Also, it does not mean that independent variables are correlated with each other, as occurs with multicollinearity. For each observation, the error term represents the distance between the actual value of the dependent variable and the predicted value. Think of the error term as the model’s “mistake.” These mistakes must not follow a pattern. If there is such a pattern, then there must be some way to improve the model so that the regression does a better job of predicting the dependent variable. A model that exhibits autocorrelation can perform better than it is.

The most common type of autocorrelation, first-order autocorrelation, is present when an observed error tends to be influenced by the observed error that immediately precedes it in the previous time period. We call this first-order autocorrelation because only one time period separates the two correlated error term observations. This can be stated as

\[ e_t = \rho e_{t-1} + u_t \]  \hspace{1cm} (7.2)

where \( e_t \) is the error from a regression in the current time period and \( e_{t-1} \) is the error from the preceding time period. The Greek letter rho, \( \rho \) (pronounced “row”), is the autocorrelation coefficient. It shows the relationship between \( e_t \) and \( e_{t-1} \) in a manner

<table>
<thead>
<tr>
<th>Year</th>
<th>Quarter</th>
<th>REVENUES</th>
<th>MARKETING</th>
<th>SUMMER</th>
<th>FALL</th>
<th>WINTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>1</td>
<td>2547.39</td>
<td>417.98</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3125.75</td>
<td>477.74</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2747.66</td>
<td>507.67</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3425.52</td>
<td>534.87</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>1</td>
<td>3168.30</td>
<td>531.38</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3559.01</td>
<td>598.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3267.48</td>
<td>601.96</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Data are available at the text website or http://www.microsoft.com/msft/download/historypivot.xls. These data were then adjusted for inflation using the urban CPI available at http://www.economagic.com or http://stats.bls.gov.
similar to how a slope coefficient works in a regression. If $\rho$ is 0.5, then on average each error will tend to be half the value of the preceding error. Any particular error $e_t$ will not be exactly half of $e_{t-1}$, the preceding error, because of the presence of $u_t$, which is the error term for Equation (7-2). $u_t$ is different from $e_t$ or $e_{t-1}$ because it follows the classical OLS assumptions. This means that $u_t$ is random and is not correlated with previous or subsequent values of $u_t$ such as $u_{t-1}$ or $u_{t+1}$. Figure 7-1 shows first-order autocorrelation, since the sign of each error term observation tends to be the same as the sign in the preceding time period.

The value of $\rho$ must fall between $-1$ and $+1$. If $\rho$ becomes larger than $+1$, following Equation (7-2), each error will tend to be larger than the one before it. As the errors get larger and larger, the regression becomes unstable; the software can no longer estimate the regression properly. When this happens, some econometricians say the regression “explodes.” If $\rho$ is less than $-1$, the errors are alternating between positive and negative, and they are getting farther and farther from zero over time until, once again, the regression explodes. If $\rho$ were to exactly equal $+1$ or $-1$, the effect of one error on the next would not die out over time. For these reasons, $\rho$ must be greater than $-1$ and less than $+1$.

If $\rho$ is zero, then one error has nothing to do with the next error, so there is no autocorrelation. The errors we see will not follow a pattern over time.

If $\rho$ is positive, the errors tend to have the same sign from one period to the next. If $e_{t-1}$ is positive, then $e_t$ tends to be positive; if $e_{t-1}$ is negative, $e_t$ tends to be negative. A positive $\rho$ indicates positive autocorrelation, also called positive serial correlation. The error term observations in Figure 7-1 exhibit positive first-order autocorrelation.

If $\rho$ is negative, the errors tend to alternate signs, indicating negative autocorrelation or negative serial correlation. In this situation, a positive observed error term is usually followed by a negative one, which is usually followed by a positive one, and so on. Figure 7-2 shows what negative autocorrelation looks like. Negative autocorrelation is less common than positive autocorrelation.

As we move from one period to the next in Figure 7-2, from top to bottom on the figure, the error term observations tend to alternate signs, going from positive to negative and back to positive again.

There are other types of autocorrelation besides first-order autocorrelation. If seasonal data are being used, the error $e_t$ could depend on the error from the same season a year ago, $e_{t-4}$. Instead of $e_t = \rho e_{t-1} + u_t$ describing the situation, as with first-order autocorrelation, $e_t = \rho e_{t-4} + u_t$ now describes the autocorrelation. This situation might occur if the dependent variable is the number of swimmers who use the outdoor community pool each quarter of the year. The error for summer might be correlated with the error from the previous summer, four seasons ago. This is more likely to occur if the regression doesn’t include an independent variable to account for the fact that more people use the pool in the summer (a dummy variable for summer).
More complex types of autocorrelation also occur. The error for the current period \( e_t \) might depend on both the error from the last period, \( e_{t-1} \), and the error from two periods ago, \( e_{t-2} \). The effect of previous errors on \( e_t \) can go back even further, causing \( e_t \) to be influenced by a whole string of previous errors. (We discuss these types of autocorrelation in Chapter 11, which focuses on time-series models.)

Autocorrelation is a problem because its presence means that useful information is missing from the model. Such information might explain the movement in the dependent variable more accurately. Like the weatherman at the beginning of this chapter, a model that exhibits autocorrelation could be doing better. The presence of autocorrelation means that a model is making a similar mistake over and over as it attempts to explain movement in the dependent variable \( Y \). Once you know that autocorrelation is present, often you can use this knowledge to improve your model.
7-2 THE SYMPTOMS

Now let’s examine the effect of autocorrelation on regression results. Using the Microsoft revenue model [Equation (7-1)] and the data from Table 7-A, we get the results in Table 7-B.

According to these OLS results, MARKETING is statistically significant at a 1% error level, and the coefficient estimate has the positive sign we would expect.\(^2\) However, autocorrelation could be affecting this result, making it appear significant when it is not. The three dummy variables are all statistically insignificant. Apparently, the market for software doesn’t have a strong seasonal component. The F-test supports this conclusion (see Section 5.3). Running the regression without the three dummy variables does not substantially change the result for MARKETING or the other statistics shown in Table 7-B.

The Gauss-Markov theorem (discussed in Chapter 2) says that ordinary least squares will have a set of desirable qualities when the classical assumptions hold. If autocorrelation is present, one of these classical assumptions does not hold. This means that some of the desirable qualities or properties described by the Gauss-Markov theorem will not be true anymore. In economics, most of the time we encounter positive

\[\text{Table 7-B} \]

Results for Microsoft Revenue Model

Dependent variable: REVENUE

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-225.88</td>
<td>75.32</td>
<td>-3.00</td>
<td>0.00</td>
</tr>
<tr>
<td>MARKETING</td>
<td>5.77</td>
<td>0.18</td>
<td>32.10</td>
<td>0.00</td>
</tr>
<tr>
<td>SUMMER</td>
<td>-11.47</td>
<td>87.60</td>
<td>0.13</td>
<td>0.90</td>
</tr>
<tr>
<td>FALL</td>
<td>64.45</td>
<td>89.39</td>
<td>0.72</td>
<td>0.47</td>
</tr>
<tr>
<td>WINTER</td>
<td>28.20</td>
<td>87.79</td>
<td>0.32</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Observations: 55
R\(^2\) = 0.95
Adjusted R\(^2\) = 0.95
Residual Sum of Squares = 2,685,801
F-statistic = 258.77
Durbin-Watson statistic = 0.44

\(^2\) The t-statistic for MARKETING has a p-value listed as 0.00. Recall from Section 3-2, Table 3-C that when this occurs, the p-value has been rounded off. In this case, the p-value is less than 0.005, and the p-value has been rounded down to 0.00.
autocorrelation, not negative autocorrelation.\(^3\) When ordinary least squares is used to estimate a model with positive autocorrelation, these three symptoms show up:

1. **Autocorrelation by itself leaves the coefficient estimates unbiased.** When autocorrelation is present, the OLS procedure still produces unbiased estimates. The expected value of an estimate \(\hat{B}\) is still the true value of \(B\). (This means that if you run the regression more than once using different data and then average all the \(\hat{B}\)'s, you should come out with the true value.) However, autocorrelation often occurs because an independent variable is missing from the model. A relevant independent variable that is missing from the model can bias the coefficient estimates of the remaining variables, even without autocorrelation. The estimates for the Microsoft revenue model may be biased, if the autocorrelation is due to a missing independent variable. There must be other important factors besides marketing and advertising expenditures that affect Microsoft's revenues. It is likely that there are independent variables missing from this model. Autocorrelation alone does not cause the estimates to be biased, but when the autocorrelation comes from a missing independent variable, the estimates will be biased.

2. **Autocorrelation increases the variance of the coefficient estimates.** The \(\hat{B}\)'s vary more from their true value than they would if there were no autocorrelation. The autocorrelation causes the \(\hat{B}\) sampling distribution to become wider. Since there is no bias, the expected value of \(\hat{B}\) is the true value. However, because autocorrelation increases the variance of the coefficient estimates, any one estimate is likely to be farther from the true value than if autocorrelation is not present. This makes the true values of the standard errors larger than they would be without autocorrelation. However, the estimates of the standard errors are smaller (symptom 3).

3. **The estimated standard errors given by ordinary least squares will be smaller than the true values.** Since these estimated standard errors end up being in the bottom half of the formula used to calculate the t-statistic, the t-statistics for the regression will be larger than they should be. When autocorrelation is present, the t-statistics will be larger than their true values. Left untreated, autocorrelation is dangerous for the researcher. Anyone examining the work can easily criticize it, pointing out that coefficients that seem to be significant may be insignificant, since the t-statistics are higher than their true values. The F-statistic, \(R^2\), and adjusted \(R^2\) may not be accurate either. These symptoms can be fatal for any model, making the results at best unclear and at worst meaningless.

### 7-3 Testing for the Illness: The Durbin-Watson Statistic

Early detection is key to curing many diseases. We can easily recognize the autocorrelation present in the Microsoft revenue model simply by looking at the pattern in Figure 7-1. Most of the time, however, autocorrelation is not as severe as this, so it is harder

\[^3\] If the autocorrelation is negative, the coefficient estimates will still be unbiased, but the effect on the slope estimate variance and the estimated standard errors is not as clear.
to recognize. Even if you think autocorrelation is present just from looking, you should back this up with objective statistical evidence. You need a test for first-order autocorrelation. The Durbin-Watson statistic provides such a test. The Durbin-Watson statistic tests for first-order autocorrelation only. Also, it does not work properly if a dependent variable from a preceding time period is used as an independent variable in the model. Most econometric software programs calculate the Durbin-Watson statistic automatically. To see how the Durbin-Watson statistic is calculated, let’s break it down by numerator and denominator. To find the numerator, follow these steps:

1. Starting with the second error term observation, find the difference between the current error term observation \( \hat{e}_2 \) and the preceding time period’s error term observation, \( \hat{e}_1 \): \( \hat{e}_2 - \hat{e}_1 \).
2. Square the number found in step 1: \( (\hat{e}_2 - \hat{e}_1)^2 \).
3. Repeat this process, advancing the time period by one each time you get back to step 1. Add the numbers from step 2 for all the time periods. The number of values you add will be one less than the total number of time periods, since you started with the second observed error term.
4. The sum you found in step 3 is the numerator in the Durbin-Watson statistic:

\[
(\hat{e}_2 - \hat{e}_1)^2 + (\hat{e}_3 - \hat{e}_2)^2 + (\hat{e}_4 - \hat{e}_3)^2 + \cdots + (\hat{e}_n - \hat{e}_{n-1})^2
\]

To find the denominator, follow these steps:

1. Take each error term observation and square it: \( \hat{e}_1^2, \hat{e}_2^2, \hat{e}_3^2, \ldots \).
2. Add all the values you found in step 1:

\[
\hat{e}_1^2 + \hat{e}_2^2 + \hat{e}_3^2 + \cdots + \hat{e}_n^2
\]

The Durbin-Watson statistic is

\[
D.W. = \frac{(\hat{e}_2 - \hat{e}_1)^2 + (\hat{e}_3 - \hat{e}_2)^2 + (\hat{e}_4 - \hat{e}_3)^2 + \cdots + (\hat{e}_n - \hat{e}_{n-1})^2}{\hat{e}_1^2 + \hat{e}_2^2 + \hat{e}_3^2 + \cdots + \hat{e}_n^2} \quad (7-3)
\]

Notice that the number of values summed in the numerator of Equation (7-3) is one less than the number of values summed in the denominator. This is how it is supposed

---

6 Using summation signs (see Section 1-4), the Durbin-Watson statistic is

\[
D.W. = \frac{\sum_{i=2}^{n}(\hat{e}_i - \hat{e}_{i-1})^2}{\sum_{i=1}^{n}\hat{e}_i^2}
\]
to work; the numerator starts with the second observed error term and the denominator starts with the first observed error term. The Durbin-Watson statistic works this way because if we start the numerator with the first observed error term instead of the second one, we would not have an initial value to subtract off.

Step 1 for the numerator takes the difference between an error term observation and the error term observation from the time period before it. This means that the order of the error term observations is important in calculating the Durbin-Watson statistic. In a time-series model, the observations are always ordered by time, but in a cross-section model, the order of observations is arbitrary. For example, a cross-sectional model using the fifty states could be arranged alphabetically. Alphabetical order has nothing to do with the characteristics of the states themselves. If we reorder the data in the cross-section so that the biggest state comes first and the smallest state last, we get a different Durbin-Watson statistic. This makes the Durbin-Watson statistic meaningless for most cross-section regressions. Fortunately, autocorrelation is not a typical problem for cross-section models, since the order of the data is arbitrary. However, autocorrelation can occur in models that use both cross-section and time-series data. (We discuss these types of models in Chapter 9.)

Changing the order of observations in a time-series regression does not eliminate autocorrelation. It just prevents the Durbin-Watson statistic from detecting it. The coefficient estimates, t-statistics, F-statistic, and adjusted $R^2$ remain the same; only the Durbin-Watson statistic changes. The autocorrelation still exists, but now it is hidden because the data are no longer in the right order. Always arrange time-series data in chronological order.

The Durbin-Watson statistic is related to the autocorrelation coefficient $\rho$. Approximately, the Durbin-Watson statistic equals $2 - 2\rho$. Some useful information can be derived from this relationship. If there is no autocorrelation, then $\rho$ is 0. This makes the Durbin-Watson statistic equal to 2. The worst possible case of first-order positive autocorrelation occurs when $\rho$ is very close to +1. If $\rho$ is equal to 1, the Durbin-Watson statistic will be $2 - 2\rho = 2 - (2 \cdot 1) = 0$. This means the closer the Durbin-Watson statistic is to zero, the more likely serious positive autocorrelation exists. For first-order negative autocorrelation, the worst possible case occurs when $\rho$ is close to −1. If $\rho$ is equal to −1, the Durbin-Watson statistic will be $2 - 2\rho = 2 - [2 \cdot (-1)] = 4$. This means that when the Durbin-Watson statistic is closer to 4, the chances of first-order negative autocorrelation increase. In summary, the Durbin-Watson statistic varies from 0 to 4: Values closer to 0 indicate positive autocorrelation; values close to 2 indicate no autocorrelation; and values closer to 4 indicate negative autocorrelation.

Most hypothesis tests use a critical value to separate the regions where the null hypothesis is rejected or not rejected. The Durbin-Watson statistic has three regions: reject the null hypothesis, do not reject the null hypothesis, and an inconclusive region. If the Durbin-Watson statistic falls in the inconclusive region, then the test is unable to reach a decision. Since there are three regions instead of the usual two, the Durbin-Watson test requires two critical values. (t-Tests and F-tests need only one critical value.) These two critical values are referred to as $d_U$ ("d-upper") and $d_L$ ("d-lower") and are given in Tables E and F in the Appendix. Use the number of independent variables in your regression to find the correct column for $d_U$ and $d_L$ in the table, and the sample size $n$ to find the correct row.
The procedure for the test differs depending on whether you are checking for positive or negative autocorrelation. When testing for positive first-order autocorrelation, use the Durbin-Watson statistic to test:

\[
H_0: \rho \leq 0 \\
H_A: \rho > 0
\]

This is a one-sided test—the null hypothesis of no autocorrelation versus the alternative hypothesis of positive autocorrelation.\(^7\)

By examining the pattern of the error term observations in Figure 7-1, we already know the Microsoft revenue model exhibits positive autocorrelation. Sometimes autocorrelation is not obvious and cannot be detected without performing the Durbin-Watson test. When the Durbin-Watson statistic is less than 2, there could be positive autocorrelation. The Durbin-Watson statistic for the Microsoft revenue model is 0.44. (See Table 7-B.) Table E in the Appendix gives Durbin-Watson \(d_L\) and \(d_U\) critical values at 5% significance for one-sided tests. There are four independent variables in the Microsoft regression, and the sample size is 55. The two critical values needed for the test are \(d_L = 1.41\) and \(d_U = 1.73.\(^8\) The decision rules for the Durbin-Watson test for positive autocorrelation are as follows:

- If the Durbin-Watson statistic is less than \(d_L\), reject the null hypothesis of no autocorrelation; assume positive autocorrelation.
- If the Durbin-Watson statistic is greater than \(d_U\), do not reject the null hypothesis of no autocorrelation; assume no autocorrelation.
- If the Durbin-Watson statistic lies between \(d_L\) and \(d_U\) (or exactly equal to either \(d_L\) or \(d_U\)), the test is inconclusive.

If the Durbin-Watson statistic is less than \(d_L\), the Durbin-Watson statistic is so far below 2 and so close to 0 that it is unlikely we would get such a low Durbin-Watson statistic when there is no positive autocorrelation. We reject the null hypothesis; there probably is positive autocorrelation. If the Durbin-Watson statistic is greater than \(d_U\), the Durbin-Watson statistic is so close to 2 that positive autocorrelation may not be present in the model. In this case, we do not reject the null hypothesis.

The Durbin-Watson statistic for the Microsoft Revenue model is below \(d_L\) (0.44 < 1.41), so we reject the null hypothesis of no autocorrelation at a 5% significance level. We can assume that there is positive autocorrelation. The results of the model must be interpreted cautiously—if they are even to be taken seriously at all, considering the problems discussed in the preceding section. In particular, because of the autocorrelation, it is not safe to say that the coefficient for MARKETING is statistically significant even though the t-test in Table 7-B reports it is significant at 1%. This t-statistic is a mistake caused by the autocorrelation. It is too high. We will discuss steps that can be taken to correct the autocorrelation problem in Sections 7-4 and 7-5.

\(^7\) Technically, the one-sided test shown here is really the null hypothesis of either no autocorrelation or negative autocorrelation versus the alternative hypothesis of positive autocorrelation.

\(^8\) These critical values are found by interpolation. For a 5% significance level, with \(k = 4\), \(d_L = 138\) for \(n = 50\) and \(d_L = 144\) for \(n = 60\). (Appendix Table E.) For \(n = 55\), we assume \(d_L = \) the average of 138 and 144, or 141. \(d_U\) is found the same way.
The decision rules for a Durbin-Watson negative autocorrelation test are different from those for positive autocorrelation. Now the null and alternative hypotheses are

\[ H_0: \rho \geq 0 \]
\[ H_A: \rho < 0 \]

When the Durbin-Watson statistic comes out greater than 2, negative autocorrelation may be present. Here are the decision rules for a Durbin-Watson test of negative autocorrelation:

- If the Durbin-Watson statistic is greater than \(4 - d_L\), reject the null hypothesis of no autocorrelation; assume negative autocorrelation.
- If the Durbin-Watson statistic is less than \(4 - d_U\), do not reject the null hypothesis of no autocorrelation; assume no autocorrelation.
- If the Durbin-Watson statistic lies between \(4 - d_L\) and \(4 - d_U\) (or exactly equal to either \(4 - d_L\) or \(4 - d_U\)), the test is inconclusive.

Since \(d_L\) will always be a smaller value than \(d_U\), \(4 - d_L\) will always be a higher value than \(4 - d_U\). If the Durbin-Watson statistic is larger than \(4 - d_L\), it is far enough from 2 and close enough to 4 that we can reject the null hypothesis of no autocorrelation. We must assume that negative autocorrelation is present. If the Durbin-Watson statistic is smaller than \(4 - d_U\), it is close enough to 2 that we do not reject the null hypothesis. Then we assume there is no autocorrelation. If the Durbin-Watson statistic is between \(4 - d_L\) and \(4 - d_U\), the test is inconclusive.

Suppose a regression with a sample size of 70 and 3 independent variables has a Durbin-Watson statistic of 2.27. Table E in the Appendix gives us \(d_L = 1.53\) and \(d_U = 1.70\) for a 5% significance test. Then \(4 - d_L = 4 - 1.53 = 2.47\), and \(4 - d_U = 4 - 1.70 = 2.30\). Since 2.27 is less than 2.30, the Durbin-Watson statistic is less than \(4 - d_U\), so we do not reject the null hypothesis. We assume there is no autocorrelation. If the Durbin-Watson statistic had been between 2.30 and 2.47, the test would be inconclusive. If the Durbin-Watson statistic had been above 2.47, we would assume there is negative autocorrelation.

7-4 TREATING THE DISEASE

There are two general approaches to dealing with autocorrelation. One approach, discussed in Section 7-5, eliminates the symptoms of autocorrelation by using an estimation method other than ordinary least squares. The best approach—discussed here—is to prevent autocorrelation from occurring in the first place, rather than just ridding the model of symptoms. Autocorrelation can infect a regression if there are relevant independent variables missing from the model. For the Microsoft revenue model, other independent variables besides marketing and advertising expenditures must be relevant. For example, Microsoft spends money on research and development (R&D) of new products. These expenditures are also likely to affect revenues. Let’s add a new independent variable RESEARCH to the model:

\[ \text{RESEARCH} = \text{Microsoft’s real expenditures on R&D, in millions of dollars} \]
It takes time for R&D expenditures to affect revenue both in theory and in practice. Therefore, we add RESEARCH to the model as a lagged variable; the value for RESEARCH will be the value from four quarters (1 year) before the current time period. This four-quarter lag is indicated by \( \text{RESEARCH}(-4) \) in Equation (7-4).

\[
\text{REVENUES} = B_0 + B_1\text{MARKETING} + B_2\text{SUMMER} + B_3\text{FALL} \\
\quad + B_4\text{WINTER} + B_5\text{RESEARCH}(-4) + e
\]  

(7-4)

The data for Microsoft’s R&D expenditures are given in Table 7-C. Data for 1986 are needed, even though the sample for the other variables starts with 1987. The first quarter value for 1986 (3.66) is necessary to have the first observation of \( \text{RESEARCH}(-4) \). The data are available at the text website or http://www.microsoft.com/msft/download/historypivot.xls. These data were then adjusted for inflation by using the urban CPI available at http://www.economagic.com or stats.bls.gov.

### Table 7-C

Data for Microsoft’s R&D Expenditures

<table>
<thead>
<tr>
<th>Year</th>
<th>First Quarter</th>
<th>Second Quarter</th>
<th>Third Quarter</th>
<th>Fourth Quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>3.66</td>
<td>4.59</td>
<td>4.55</td>
<td>6.34</td>
</tr>
<tr>
<td>1987</td>
<td>6.27</td>
<td>7.07</td>
<td>9.62</td>
<td>10.40</td>
</tr>
<tr>
<td>1988</td>
<td>12.92</td>
<td>13.61</td>
<td>15.11</td>
<td>17.45</td>
</tr>
<tr>
<td>1989</td>
<td>18.90</td>
<td>20.22</td>
<td>23.26</td>
<td>26.22</td>
</tr>
<tr>
<td>1990</td>
<td>31.24</td>
<td>31.70</td>
<td>35.72</td>
<td>39.27</td>
</tr>
<tr>
<td>1991</td>
<td>37.83</td>
<td>39.82</td>
<td>43.90</td>
<td>50.84</td>
</tr>
<tr>
<td>1992</td>
<td>53.37</td>
<td>61.52</td>
<td>63.88</td>
<td>71.88</td>
</tr>
<tr>
<td>1993</td>
<td>73.38</td>
<td>76.98</td>
<td>80.13</td>
<td>94.67</td>
</tr>
<tr>
<td>1994</td>
<td>91.34</td>
<td>101.60</td>
<td>104.74</td>
<td>113.61</td>
</tr>
<tr>
<td>1995</td>
<td>117.99</td>
<td>130.75</td>
<td>143.26</td>
<td>171.88</td>
</tr>
<tr>
<td>1996</td>
<td>179.35</td>
<td>194.85</td>
<td>214.15</td>
<td>256.15</td>
</tr>
<tr>
<td>1997</td>
<td>261.33</td>
<td>280.27</td>
<td>294.72</td>
<td>323.91</td>
</tr>
<tr>
<td>1998</td>
<td>372.45</td>
<td>398.12</td>
<td>373.93</td>
<td>450.70</td>
</tr>
<tr>
<td>1999</td>
<td>395.50</td>
<td>430.20</td>
<td>397.05</td>
<td>558.64</td>
</tr>
<tr>
<td>2000</td>
<td>490.78</td>
<td>530.47</td>
<td>571.92</td>
<td></td>
</tr>
</tbody>
</table>
when the regression starts in the first quarter of 1987. The regression’s last observation
is the third quarter of 2000; the value of $\text{RESEARCH}(-4)$ for this observation is taken
from the third quarter of 1999. It is 397.05. Most econometrics software programs have
features that allow you to use lagged independent variables without any fuss. (More
will be said about lagged independent variables in Chapter 11.)

Adding $\text{RESEARCH}(-4)$ may reduce or eliminate the autocorrelation problem, if
it is an appropriate independent variable for this model. Figure 7-3 shows the error term
observations we get when we include $\text{RESEARCH}(-4)$ in the model (using the data
from Tables 7-A and 7-C).

It is not clear from looking at Figure 7-3 whether some positive autocorrelation is
still present or whether it has been eliminated. Certainly, if some positive autocorrela-
tion remains, it is not as pronounced as before. To see this, compare parts a and b (Fig-
ure 7-1 has been repeated as part b). The error term observations change signs (move
across the center line) much more frequently in part a than they do in part b. In part a
it is a lot harder to predict the sign of an observed error simply by noting whether the
preceeding term is positive or negative, so we cannot tell whether autocorrelation is pres-
ent just by looking. The Durbin-Watson test may help.

Table 7-D presents the results for the Microsoft revenue model with $\text{RESEARCH}(-4)$
included. If $\text{RESEARCH}$ is lagged by a different number of quarters (instead of 4), the
results are similar. The Durbin-Watson statistic here is 2.14, a lot closer to 2 than the
Durbin-Watson statistic for the original model, 0.44 (see Table 7-B). The number of
independent variables in the model has changed because we added $\text{RESEARCH}$, so we
must use Table E in the Appendix again to find the appropriate $d_L$ and $d_U$. Now $d_L = 1.38$ and $d_U = 1.77$ at 5% significance with 5 independent variables and 55 observa-
tions.\(^9\) The Durbin-Watson statistic is greater than $d_U$ since 2.14 > 1.77. This test result
allows us to assume that positive autocorrelation is no longer present.\(^10\)

There are probably other relevant independent variables missing from this model.
It is unlikely that $\text{MARKETING}$ and $\text{RESEARCH}(-4)$ are the only important inde-
pendent variables.\(^11\) Also, a single-equation model may not describe Microsoft’s reve-
 nue generating process very well. (Multiple-equation regression models are discussed
in Chapter 10.) The important point here is that adding a missing relevant independent
variable cured our autocorrelation problem.

When a relevant independent variable is left out, the explanatory power it brings to
the model is assigned elsewhere by the OLS procedure. When $\text{RESEARCH}(-4)$ was
omitted, part of its explanatory power was assigned to the error term. This caused the

\(^9\) Once again, these critical values are found by interpolating between the critical values for $n = 50$ and $n
= 60$ (see last footnote).

\(^10\) It may seem obvious there is no positive autocorrelation since the Durbin-Watson statistic is greater than
2. However, if you look at Table E in the Appendix for low sample sizes, there are some $d_U$ values that are
slightly greater than 2. This means that even if the Durbin-Watson statistic is greater than 2, it is still possi-
ble for a Durbin-Watson positive autocorrelation test to be inconclusive.

\(^11\) $\text{MARKETING}$ and $\text{RESEARCH}$ have a high correlation coefficient, indicating multicolli-
nearity. Since both variables are highly significant, as discussed in Section 6-4, it is reasonable to use both variables in this
regression.
Figure 7-3
Error Term Observations from Microsoft Revenues Model with RESEARCH(−4) Included.

(a) With RESEARCH(−4)  (b) Without RESEARCH(−4) (original model)
error term to be correlated with RESEARCH(−4). If RESEARCH(−4) is correlated with itself over time, which is likely, this causes the error term to be correlated with itself over time, giving us autocorrelation. MARKETING will also be given some of the credit for explaining REVENUE that should have been credited to RESEARCH(−4). This will cause the coefficient estimate for MARKETING to be biased. The higher the correlation between MARKETING and the missing RESEARCH(−4), the more biased the coefficient estimate for MARKETING. As discussed in Section 7-2, autocorrelation by itself does not produce biased estimates. However, a missing independent variable does produce biased estimates, so regressions with autocorrelation caused by missing relevant independent variables will have biased estimates.

Autocorrelation can also be present in a model if the functional form of the regression is incorrect. Suppose we take the natural logs of REVENUE, MARKETING, and RESEARCH(−4) before we estimate the regression in the Microsoft revenues model.

\[
\ln(\text{REVENUES}) = B_0 + B_1 \ln(\text{MARKETING}) + B_2 \text{SUMMER} + B_3 \text{FALL} \\
+ B_4 \text{WINTER} + B_5 \ln(\text{RESEARCH(−4)}) + \epsilon
\]  

(7-5)

When we estimate this model using the same data as before, the autocorrelation returns, even though we have included RESEARCH(−4). The Durbin-Watson statistic is 0.76, indicating positive autocorrelation at a 5% (and even a 1%) significance level. The linear model containing RESEARCH(−4) gives better results since it does not exhibit autocorrelation.
The best approach to the autocorrelation problem is to try to formulate the model correctly in the first place. Here, omitting the R&D variable from the model caused positive first-order autocorrelation. This is one reason why it is important to think about theory when formulating models. For regressions dealing with economics, economic theory can guide you in including relevant independent variables and in using the correct functional form for the model. Whatever your field, using theory to formulate the model correctly to begin with prevents a lot of trouble.

7-5 TREATING THE SYMPTOMS

There are situations in which autocorrelation cannot be eliminated by changing the model. Perhaps the omitted independent variables that would help are not available or cannot be measured. Rather than toss the model out we treat the symptoms. This is a second-best solution because we have not cured the disease. It is always preferable to use the right model so that there is no autocorrelation. But, when this can’t be done, “brute force” methods are available to make the error term observations uncorrelated, so that the symptoms of autocorrelation do not occur. (The estimates will be efficient and the t-statistics will be correct, and so on.) The regression will appear normal—will not exhibit autocorrelation—but the underlying model is the one that produced the autocorrelation in the first place.

Consider a generic simple regression model:

\[ Y_t = B_0 + B_1X_t + e_t \quad (7-6) \]

The t subscripts indicate that this is a time-series regression and that for each observation the values of Y and X are both from the same time period.\(^\text{12}\) If this model has first-order autocorrelation, then, as expressed by Equation (7-2), this period’s error and the preceding one are related: \( e_t = \rho e_{t-1} + u_t \). Remember, \( u_t \) represents an error term that follows OLS assumptions, so \( u_t \) is independent of its preceding value, \( u_{t-1} \). We can force the autocorrelation out of the model if we can replace \( e_t \) by \( u_t \), since \( u_t \) meets the OLS assumptions. Here is how this is done.

First we rearrange Equation (7-2) by subtracting \( \rho e_{t-1} \) from each side to get

\[ e_t - \rho e_{t-1} = u_t \quad (7-7) \]

We want to find a way to eliminate the correlated errors (the e’s) from the regression and replace them with random errors (the u’s). If we can rearrange the regression so that \( e_t - \rho e_{t-1} \) appears in the regression, then we can use Equation (7-7) to substitute the good error term (u’s) for the bad ones (e’s).

First, consider Equation (7-6). This regression is not just for time period t but for all the time periods included in the data. Instead of using \( Y_t = B_0 + B_1X_t + e_t \), it is valid to move everything back one period. Then we have

\[ Y_{t-1} = B_0 + B_1X_{t-1} + e_{t-1} \quad (7-8) \]

\(^{12}\) Throughout most of the book, the t subscripts have been omitted for simplicity, but here they are necessary.
Next, multiply both sides of (7-8) by the autocorrelation coefficient $\rho$:

$$\rho Y_{t-1} = \rho B_0 + \rho B_1 X_{t-1} + \rho e_{t-1} \quad (7-9)$$

We are almost there. We want to get $e_t - \rho e_{t-1}$ into the regression so we can use (7-7) to replace $e_t - \rho e_{t-1}$ with $u_t$, forcing out the autocorrelation. Take the original regression (7-6) and subtract (7-9) from it, on both the left- and right-hand sides of the equation:

$$Y_t = B_0 + B_1 X_t + e_t$$

$$- \rho Y_{t-1} = -(\rho B_0 + \rho B_1 X_{t-1} + \rho e_{t-1})$$

$$Y_t - \rho Y_{t-1} = B_0 - \rho B_0 + (B_1 X_t - \rho B_1 X_{t-1}) + e_t - \rho e_{t-1}$$

The bottom line, the answer to this subtraction, can be rewritten as

$$Y_t - \rho Y_{t-1} = B_0(1 - \rho) + B_1 (X_t - \rho X_{t-1}) + e_t - \rho e_{t-1} \quad (7-10)$$

Now $e_t - \rho e_{t-1}$ appears in the regression, and we can eliminate the $e$ terms that exhibit autocorrelation by substituting $u_t$ in for $e_t - \rho e_{t-1}$ [Equation (7-7)]:

$$Y_t - \rho Y_{t-1} = B_0(1 - \rho) + B_1 (X_t - \rho X_{t-1}) + u_t \quad (7-11)$$

Equation (7-11) does not have autocorrelation. Think of $Y_t - \rho Y_{t-1}$ as one variable, the new dependent variable. $B_0(1 - \rho)$ is the new intercept and $X_t - \rho X_{t-1}$ is the new independent variable. This is called a generalized difference equation because we find it by taking the difference of the two equations. This process also works with more than one independent variable; we used only one here for simplicity. The generalized difference equation is a type of generalized least squares, a different estimation method from ordinary least squares.

We need an estimate for the autocorrelation coefficient $\rho$ before we can estimate the regression in Equation (7-11). Both the Cochrane-Orcutt and AR(1) methods find an estimate of $\rho$ and complete the regression estimation. The Cochrane-Orcutt method has been the most commonly used method, but now that econometrics software is more sophisticated, the AR(1) method is becoming more popular.

### The Cochrane-Orcutt Method

This method estimates $\rho$ several times until the estimates stop changing, or converge. The Cochrane-Orcutt method is often called an iterative process because it repeats certain steps over and over.\(^{13}\) The steps are:

1. Estimate the regression using ordinary least squares.
2. Use the error term observations from step 1 to estimate $e_t = \rho e_{t-1} + u_t$, getting an estimate for $\rho$.

3. Use the estimate for $\rho$ along with the data for the dependent and independent variables to estimate the generalized difference equation, (7-11).

4. Using the error term observations from step 3, go back to step 2 and estimate $\rho$ again. Repeat this process until the estimate of $\rho$ stays about the same; then estimate the generalized difference equation one last time.

Your results will include an estimate for $\rho$, along with the coefficient estimates, t-statistics, and other statistics you expect. Most software programs that provide the Cochrane-Orcutt method also show you the intermediate estimates for $\rho$ that were used during the process. The sample size will be one less than the number of observations available, since no observation precedes the first one to use as $Y_{t-1}$ or $X_{t-1}$ in Equation (7-11). The first observation is automatically omitted, because it provides $Y_{t-1}$ and $X_{t-1}$ for the second observation, which starts the regression. The coefficient estimate results given by the Cochrane-Orcutt process are usually very close to the results given by the AR(1) method. Now let's look at an example of AR(1) results.

**The AR(1) Method**

The AR(1) method uses a nonlinear technique to estimate the autocorrelation coefficient $\rho$ and the B coefficient estimates all at once. The details of the method are complex and will not be covered here. AR(1) is considered a more powerful method, but for practical purposes, the differences between the AR(1) and Cochrane-Orcutt methods seem small.

Suppose Microsoft released data only for revenue and marketing, but not for R&D. Then we might resort to using Cochrane-Orcutt or AR(1) to deal with the autocorrelation problem, since RESEARCH$(-4)$ cannot be used as an independent variable. Figure 7-4 shows the error term observations when the AR(1) process is applied to the original Microsoft revenue model, which exhibited positive first-order autocorrelation [RESEARCH$(-4)$ is not included]. These error term observations are observations of $u_t$ from Equation (7-11). They should follow the OLS assumptions and should not exhibit autocorrelation.

As shown in Figure 7-4a, the error term observations do not exhibit a clear pattern after we apply AR(1) to the regression. Table 7-E presents our results using the AR(1) method. *(Note: The variable listed as AR(1) is not an independent variable at all. It is the AR(1) method’s estimate of $\rho$, the autocorrelation coefficient.)*

The variables listed in Table 7-E are actually generalized difference variables. This means that where Table 7-E lists the variable MARKETING, it really means the variable MARKETING$\_t - \rho$MARKETING$\_{t-1}$, matching the format given in Equation (7-11). The same is true with the dependent variable. That is, the dependent variable here is not REVENUE, it is REVENUE$\_t - \rho$REVENUE$\_{t-1}$. Each observation in the time series is now a term containing values from the current time period and the preceding period.

The estimate of $\rho$ is 0.80, which lies between 0 and 1, as we expect for a $\rho$ when positive autocorrelation is present. MARKETING is still statistically significant at a 1% significance level, so the original autocorrelation did not cause us to mistake an insignificant variable for a significant one. Note that the t-statistic for the MARKETING coefficient fell from 32.10 to 11.23 (see Table 7-B). This supports the idea that autocorrelation inflates t-statistics. The F-test shows that as a group the dummy vari-
Figure 7-4
Comparison of Error Term Observations Using AR(1) for Microsoft Revenue Model.

(a) AR(1) forces autocorrelation out

(b) Autocorrelation in original model
ables are not statistically significant at a 5% error level. Removing the dummy variables from the regression using the AR(1) method does not cause much change in the remaining coefficient estimates or statistics.

Both the Cochrane-Orcutt and AR(1) methods force first-order autocorrelation out of models, eliminating the symptoms or problems that accompany autocorrelation. It is always better to cure autocorrelation in your model by identifying what is causing the autocorrelation in the first place. Use the brute force methods of Cochrane-Orcutt and AR(1) only as a last resort. Otherwise, you may overlook an important variable that adds insight to your research.

### SUMMARY

1. **Autocorrelation** (also called **serial correlation**) occurs when the errors follow a pattern; it is a violation of the classical assumption that the errors are independent of each other. Autocorrelation is a common problem in time-series regressions. It indicates that a relevant independent variable is missing from the model.

2. **First-order autocorrelation** is present when an error is influenced by the error from the preceding time period. The autocorrelation coefficient $\rho$ (Greek letter rho, pronounced “row”) shows the relationship between $e_t$ and $e_{t-1}$ in a manner

### Table 7-E

Results for Microsoft Revenue Model Estimated by AR(1) Excluding RESEARCH(−4)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−151.27</td>
<td>171.56</td>
<td>−0.88</td>
<td>0.38</td>
</tr>
<tr>
<td>MARKETING</td>
<td>5.45</td>
<td>0.49</td>
<td>11.23</td>
<td>0.00</td>
</tr>
<tr>
<td>SUMMER</td>
<td>−10.50</td>
<td>37.83</td>
<td>−0.28</td>
<td>0.78</td>
</tr>
<tr>
<td>FALL</td>
<td>65.99</td>
<td>45.01</td>
<td>1.47</td>
<td>0.15</td>
</tr>
<tr>
<td>WINTER</td>
<td>17.22</td>
<td>39.07</td>
<td>0.44</td>
<td>0.66</td>
</tr>
<tr>
<td>AR(1) [estimate of $\rho$]</td>
<td>0.80</td>
<td>0.090</td>
<td>8.9</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Observations: 54  
$R^2 = 0.98$  
Adjusted $R^2 = 0.98$  
Residual Sum of Squares = 1,027,335  
F-statistic = 542.99  
Durbin-Watson statistic = 2.60
similar to how a slope coefficient works in a regression. The value of $\rho$ must fall between $-1$ and $+1$.

3. **Positive autocorrelation** occurs when the error term tends to keep the same sign from one period to the next. **Negative autocorrelation** occurs when the sign of the error term tends to alternate back and forth between positive and negative. Positive autocorrelation is more common than negative autocorrelation in economics.

4. There are other types of autocorrelation besides first-order autocorrelation. For example, if seasonal data are used, the error $e_t$ could depend on the error from the same season a year ago, $e_{t-4}$.

5. Autocorrelation:
   - leaves the coefficient estimates unbiased (unless a relevant independent variable is missing)
   - increases the variance of the coefficient estimates
   - makes the estimated standard errors given by ordinary least squares too small

Autocorrelation is a serious problem because it makes the t-statistics larger than they should be. The misleadingly large t-statistics can make you think that a slope estimate is statistically significant when it is not.

6. The **Durbin-Watson statistic** can be used to test for first-order autocorrelation, if the dependent variable does not also appear as a lagged independent variable in the model. The Durbin-Watson ranges from 0 to 4. The closer the Durbin-Watson statistic is to 0, the more likely positive first-order autocorrelation is present. The closer it is to 2, the more likely there is no autocorrelation. The closer it is to 4, the more likely negative first-order autocorrelation is present.

7. Different approaches are used to treat autocorrelation. The best approach is to reformulate the model to eliminate the autocorrelation. Sometimes this is accomplished by adding a missing independent variable or by changing the functional form of the regression. However, there are situations in which autocorrelation cannot be eliminated by changing the model. In these cases, a generalized difference equation is used.

$Y_t - \rho Y_{t-1} = B_0(1 - \rho) + B_1(X_t - \rho X_{t-1}) + u_t \quad (7-11)$

An estimate of $\rho$ must be found to estimate a generalized difference equation. Either the **Cochrane-Orcutt method** or the **AR(1) method** can be used to estimate $\rho$ and the generalized difference equation. The Cochrane-Orcutt method estimates $\rho$ and then estimates Equation (7-11), repeating the process until the estimate of $\rho$ stays about the same. The AR(1) method is a nonlinear method that estimates $\rho$ and the other coefficients all at once. The coefficient estimate results given by the two methods are usually close to each other.

**EXERCISES**

1. Try to explain these terms without looking them up.
   - serial correlation
   - first-order autocorrelation
• positive autocorrelation
• negative autocorrelation
• Durbin-Watson statistic
• Generalized difference equation
• Cochrane-Orcutt method
• AR(1)

2. Autocorrelation causes problems for ordinary least squares regression. Explain the autocorrelation problem in a simple way as you would to a friend who doesn’t know much about econometrics.

3. Autocorrelation, if left unchecked, can be particularly dangerous for a researcher who has found statistically significant results. Explain why this is so.

4. a. Besides autocorrelation, what is the biggest problem with the Microsoft revenues model as expressed by Equation (7-1)?
   b. Is your answer to part a somehow related to the autocorrelation problem? Explain.

5. Conduct a Durbin-Watson test at a 5% error level for each case given. Is the null hypothesis of no first-order autocorrelation rejected, not rejected, or is the test inconclusive? Use the appropriate one-sided test for each Durbin-Watson statistic given.
   a. D.W. = 1.27, N = 40, k = 2
   b. D.W. = 1.27, N = 25, k = 2
   c. D.W. = 2.45, N = 80, k = 5
   d. D.W. = 2.45, N = 80, k = 1

6. a. In Exercise 5, what does the difference between your answers to parts a and b tell you about how the Durbin-Watson test works?
   b. What does the difference between your answers to parts c and d in Exercise 5 tell you about how the Durbin-Watson test works?

7. The text indicates that when we add missing independent variables we are treating the illness of autocorrelation, but if we use a generalized difference equation we are just treating its symptoms. What does this mean? In other words, what is the distinction between adding missing variables to treat the illness and using a generalized difference equation to treat the symptoms?

8. The Cochrane-Orcutt method can be used to estimate a generalized least squares equation when autocorrelation is present. Without looking at the text or your notes, explain how the procedure works, including how the equation is estimated.

9. Using the data from Table 7-A, use ordinary least squares to estimate the Microsoft revenue model again for the first three years only (1987, 1988, 1989), so that the sample size is 12. Below are the same data for the 12 quarters of 1987–1989, but the order of the data is different. Run the regression again using the order shown here. Are the Durbin-Watson statistics the same?
10. What does the Durbin-Watson statistic mean for a cross-section regression? (*Hint:* Consider your answer to Exercise 9.)

11. Combine the data in Tables 2-A and 2-C to create one data set for the DVD expenditures model. (Make sure you keep the data in order, so that the data from Table 2-A represents the first year of data, and that the data in Table 2-C follows, representing the second year. Keep the months in order, also.) Use this combined data set to estimate the DVD expenditures model [see Equation (2-2)]. Conduct a Durbin-Watson test for first-order autocorrelation using a 5% error level.

12. Use the AR(1) method to correct for first-order autocorrelation in Exercise 11. Compare the AR(1) results to the original results you found in Exercise 11. What does your answer to the preceding exercise have to do with this comparison of results?

13. You may see a time-series regression that contains a trend variable, sometimes called T. The trend variable just increases one unit for every unit of time that passes in the data. To set up a trend variable for the Microsoft revenues model, we would use \( T = 1 \) for the first quarter, 2 for the second quarter, 3 for the third quarter, and so forth, until the last quarter of data where \( T = 55 \) (since \( n = 55 \)). The trend variable is supposed to represent any linear trend over time that is not being captured by other variables in the model. However, this is cheating in a way, because although you may increase the explanatory power of your model by including the trend variable, you also know that the trend variable really just represents some unknown missing variables. A significant trend variable means that

<table>
<thead>
<tr>
<th>Year</th>
<th>Quarter</th>
<th>REVENUES</th>
<th>MARKETING</th>
<th>SUMMER</th>
<th>FALL</th>
<th>WINTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1987</td>
<td>1</td>
<td>60.02</td>
<td>13.44</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>86.68</td>
<td>22.54</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>85.66</td>
<td>18.36</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>71.62</td>
<td>16.80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1988</td>
<td>1</td>
<td>88.74</td>
<td>23.26</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>141.27</td>
<td>40.72</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>136.02</td>
<td>32.75</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>132.73</td>
<td>31.48</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1989</td>
<td>1</td>
<td>145.48</td>
<td>37.81</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>175.58</td>
<td>45.29</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>158.02</td>
<td>40.91</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>169.81</td>
<td>46.90</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
a linear trend is associated with changes in the dependent variable, but this doesn’t
tell you what is driving the trend or what is missing from the model.
a. Estimate the following regression using the data in Table 7-A.

\[
\text{REVENUES} = B_0 + B_1 \text{MARKETING} + B_2 \text{SUMMER} + B_3 \text{FALL} \\
+ B_4 \text{WINTER} + B_5 T + e
\]

where \( T = 1 \) for the first observation, \( 2 \) for the second observation, and so on, so that the last observation has \( T = 55 \).
b. Is the slope estimate for \( T \) statistically significant at a 5% error level? If so, does this add to our understanding of Microsoft revenues?
c. The original Durbin-Watson statistic (without \( T \) in the model) indicated auto-
correlation (see Table 7-B). Does the inclusion of \( T \) in the model change the results of the Durbin-Watson test?