EE325

AUTOCORRELATION: WHAT HAPPENS IF THE ERROR TERMS ARE CORRELATED?

The powerpoint is based on Basic Econometrics, Damodar N. Gujarati and Dawn C. Porter., McGraw Hill, 2009.

- **In this chapter we take a critical look at the following questions:**
- **1. What is the nature of autocorrelation?**
- **2. What are the theoretical and practical consequences of autocorrelation?**
- **3. Since the assumption of no autocorrelation relates to the unobservable disturbances u_t, *how does one know that there is autocorrelation in any* given situation? Notice that we now use the subscript *t to emphasize that we* are dealing with time series data.**
- **4. How does one remedy the problem of autocorrelation?**

- **THE NATURE OF THE PROBLEM**
- **Autocorrelation may be defined as "*correlation between members of series of observations ordered in time* [as in time series data] or *space* [as in cross-sectional data].'' the CLRM assumes that:**
- $E(u_i u_j) = 0 \qquad i \neq j \qquad (3.2.5)$
- **Put simply, the classical model assumes that the disturbance term relating to any observation is not influenced by the disturbance term relating to any other observation.**
- **For example, if we are dealing with quarterly time series data involving the regression of *output on labor and capital* inputs and if, say, there is a labor strike affecting output in one quarter, there is no reason to believe that this disruption will be carried over to the next quarter. That is, if output is lower this quarter, there is no reason to expect it to be lower next quarter. Similarly, if we are dealing with cross-sectional data involving the regression of *family consumption* expenditure on family income, the effect of an increase of one family's income on its consumption expenditure is not expected to affect the consumption expenditure of another family.**

- **However, if there is autocorrelation,**
- $E(u_i u_j) \neq 0 \qquad i \neq j \qquad (12.1.1)$
- **In this situation, the disruption caused by a strike this quarter may very well affect output next quarter, or the increases in the consumption expenditure of one family may very well prompt another family to increase its consumption expenditure.**
- **In Figure 12.1. Figure 12.1*a to d shows that there* is a discernible pattern among the *u's. Figure 12.1a shows a cyclical pattern;* Figure 12.1*b and c suggests an upward or downward linear trend in the disturbances;* whereas Figure 12.1*d indicates that both linear and quadratic* trend terms are present in the disturbances. Only *Figure 12.1e indicates no systematic pattern*, supporting the non-autocorrelation assumption of the classical linear regression model.**

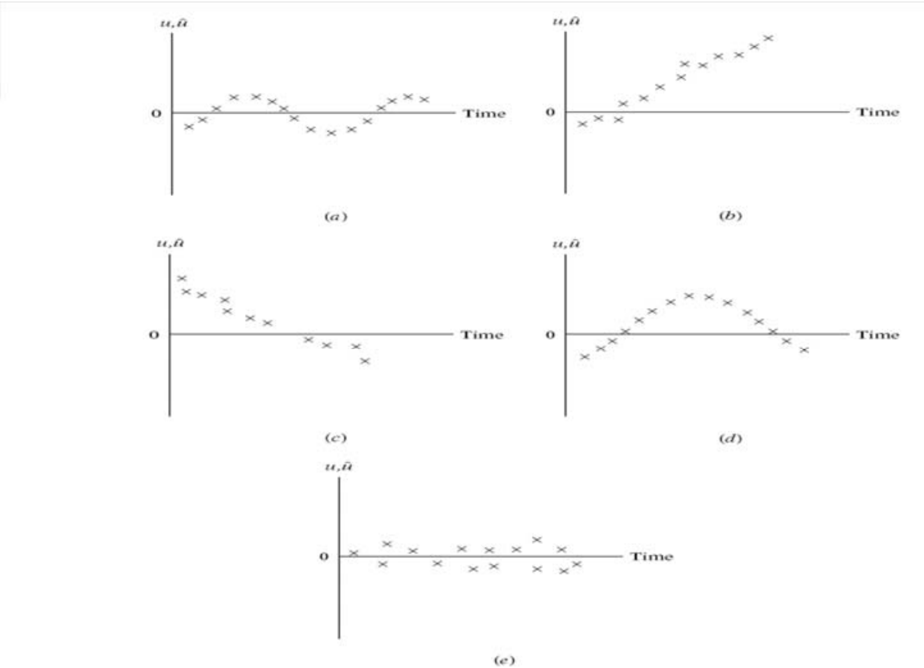

FIGURE 12.1 Patterns of autocorrelation and nonautocorrelation.

- *Why does serial correlation occur*? There are several reasons:

1. Inertia. As is well known, time series such as GNP, price indexes, production, employment, and unemployment exhibit (business) cycles. Starting at the bottom of the recession, when economic recovery starts, most of these series start moving upward. In this upswing, the value of a series at one point in time is greater than its previous value. Thus there is a momentum'' built into them, and it continues until something happens (e.g., increase in interest rate or taxes or both) to slow them down.

2. Specification Bias: *Excluded Variables Case*. In empirical analysis the researcher often starts with a plausible regression model that may not be the most "perfect'' one. For example, the researcher may plot the residuals $\hat{u}_i$ *obtained from the fitted regression and may observe patterns* such as those shown in Figure 12.1*a to d. These residuals (which are proxies* for u_i*) may suggest that some variables that were originally candidates but* were not included in the model for a variety of reasons should be included. Often the inclusion of such variables removes the correlation pattern observed among the residuals.

- **For example, suppose we have the following demand model:**
- $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + u_t$ *(12.1.2)*
- **where *Y = quantity of beef demanded,* X_2 *= price of beef,* X_3 *= consumer***
- **income, X_4 *= price of poultry, and t = time. However, for some reason we run***
- **the following regression:**
- $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + v_t$ *(12.1.3)*
- **Now if (12.1.2) is the "correct'' model or the "truth'' or true relation, running (12.1.3) is tantamount to letting $v_t = \beta_4 X_{4t} + u_t$. *And to the extent the***
- **price of *poultry* affects the consumption of beef, the error or disturbance term *v will reflect a systematic pattern, thus creating (false) autocorrelation.***
- **A simple test of this would be to run both (12.1.2) and (12.1.3) and see whether autocorrelation, if any, observed in model (12.1.3) disappears when (12.1.2) is run.**

3. Specification Bias*: Incorrect Functional Form*. Suppose the "true'' or correct model in a cost-output study is as follows:

- *Marginal* $cost_i = \beta_1 + \beta_2\, output_i + \beta_3\, output^2_i + u_i$ *(12.1.4)*
- **but we fit the following model:**
- *Marginal* $cost_i = \alpha_1 + \alpha_2\, output_i + v_i$ *(12.1.5)*
- **The marginal cost curve corresponding to the "true'' model is shown in Figure 12.2 along with the "*incorrect*'' linear cost curve.**
- **As Figure 12.2 shows, between points *A and B the linear marginal cost* curve will consistently overestimate the true marginal cost, whereas beyond these points it will consistently underestimate the true marginal cost. This result is to be expected, because the disturbance term v_i *is, in fact, equal to* $output^2 + u_i$*, and hence will catch the systematic effect of the* $output^2$ *term* on marginal cost. In this case, v_i *will reflect autocorrelation because of the* use of an incorrect functional form.**

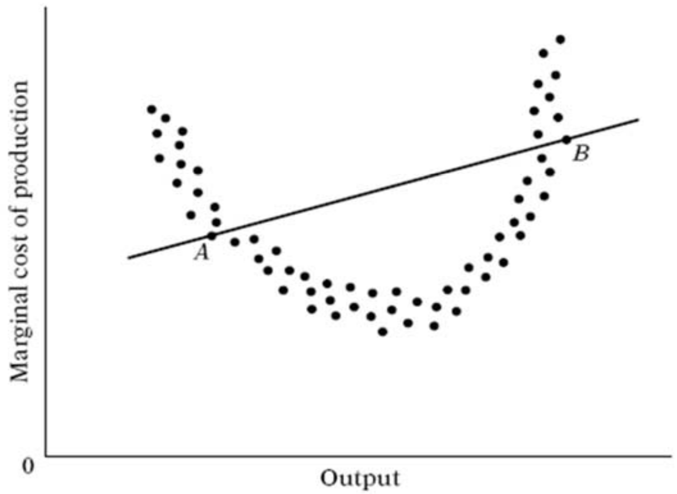

FIGURE 12.2 Specification bias: incorrect functional form.

4. Cobweb Phenomenon. The supply of many agricultural commodities reflects the so-called cobweb phenomenon, where supply reacts to price with a lag of one time period because supply decisions take time to implement (the gestation period). Thus, at the beginning of this year's planting of crops, farmers are influenced by the price prevailing last year, so that their supply function is

- **Supply$_t$ = $\beta_1 + \beta_2 P_{t-1} + u_t$ *(12.1.6)***
- **Suppose at the end of period *t, price P_t turns out to be lower than P_{t-1}.* Therefore, in *period$_{t+1}$ farmers may very well decide to produce less than* they did in *period t. Obviously, in this situation the disturbances u_t are not* expected to be random because if the farmers overproduce in *year t, they* are likely to reduce their *production in t + 1, and so on, leading to a Cobweb* pattern.**

5. Lags. In a time series regression of consumption expenditure on income, it is not uncommon to find that the consumption expenditure in the current period depends, among other things, on the consumption expenditure of the previous period. That is,

- **Consumption$_t$ = $\beta_1 + \beta_2$ *income$_t$* + β_3 *consumption$_{t-1}$* + u_t *(12.1.7)***
- **A regression such as (12.1.7) is known as autoregression because one of the explanatory variables is the lagged value of the dependent variable. The rationale for a model such as (12.1.7) is simple. Consumers do not change their consumption habits readily for *psychological, technological, or institutional* reasons. Now if we neglect the lagged term in (12.1.7), the resulting error term will reflect a systematic pattern due to the influence of lagged consumption on current consumption.**

6. "Manipulation'' of Data. In empirical analysis, the raw data are often "manipulated.'' For example, in time series regressions involving *quarterly data*, such data are usually derived from the monthly data by simply adding three monthly observations and dividing the sum by 3. This averaging introduces smoothness into the data by dampening the fluctuations in the monthly data. Therefore, the graph plotting the quarterly data looks much smoother than the monthly data, and this smoothness may itself lend to a systematic pattern in the disturbances, thereby introducing autocorrelation.

- **Another source of manipulation is *interpolation* or *extrapolation* of data. For example, the Census of Population is conducted every 10 years in this country, the last being in 2000 and the one before that in 1990. Now if there is a need to obtain data for some year within the intercensus period 1990–2000, the common practice is to interpolate on the basis of some ad hoc assumptions. All such data "massaging'' techniques might impose upon the data a systematic pattern that might not exist in the original data.**

7. Data Transformation. consider the following model:

- $Y_t = \beta_1 + \beta_2 X_t + u_t$ *(12.1.8)*
- where, say, *Y = consumption expenditure and X = income. Since (12.1.8)* holds true at every time period, it holds true also in the previous time period, ($t - 1$)*. So, we can write (12.1.8) as:*
- $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$ *(12.1.9)*
- Y_{t-1}, X_{t-1}, *and* u_{t-1} *are known as the lagged values of Y, X, and u, respectively,* here lagged by one period. Now if we subtract (12.1.9) from (12.1.8), we obtain
- $\Delta Y_t = \beta_2 \Delta X_t + \Delta u_t$ *(12.1.10)*
- where Δ, known as the first difference operator, tells us to take successive differences of the variables in question. Thus, $Y_t = (Y_t - Y_{t-1})$, $X_t = (X_t - X_{t-1})$, *and* $u_t = (u_t - u_{t-1})$*. For empirical purposes, we write* (12.1.10) as
- $Y_t = \beta_2 X_t + v_t$ *(12.1.11)*
- where $v_t = u_t = (u_t - u_{t-1})$.

- Equation (12.1.9) is known as the *level form* and Eq. (12.1.10) is known as *the (first) difference form*. Both forms are often used in empirical analysis. For example, if in (12.1.9) *Y and X represent the logarithms of* consumption expenditure and income, then in (12.1.10) *Y and X will* represent changes in the logs of consumption expenditure and income. But as we know, a change in the log of a variable is a relative change, or a percentage change, if the former is multiplied by 100. So, instead of studying relationships between variables in the level form, we may be interested in their relationships in the growth form.
- Now if the error term in (12.1.8) satisfies the standard OLS assumptions, particularly the assumption of no autocorrelation, it can be shown that the error term v_t *in (12.1.11) is autocorrelated.* It may be noted here that models like (12.1.11) are known as dynamic regression models, that is, models involving lagged regressands.
- The point of the preceding example is that sometimes autocorrelation may be induced as a result of transforming the original model.

8. Nonstationarity. We mentioned in Chapter 1 that, while dealing with time series data, we may have to find out if a given time series is stationary.

- a time series is stationary if its characteristics *<u>(e.g., mean, variance, and covariance)</u>* are *time invariant; that is, they do not change over* time. If that is not the case, we have a nonstationary time series. In a regression model such as
- $Y_t = \beta_1 + \beta_2 X_t + u_t$ *(12.1.8)*
- it is quite possible that both *Y and X are nonstationary and therefore the error u is also nonstationary.* In that case, the error term will exhibit autocorrelation.
- It should be noted also that autocorrelation can be positive (Figure 12.3*a)* as well as negative, although most economic time series generally exhibit positive autocorrelation because most of them either move upward or downward over extended time periods and do not exhibit a constant up-and-down movement such as that shown in Figure 12.3*b*.

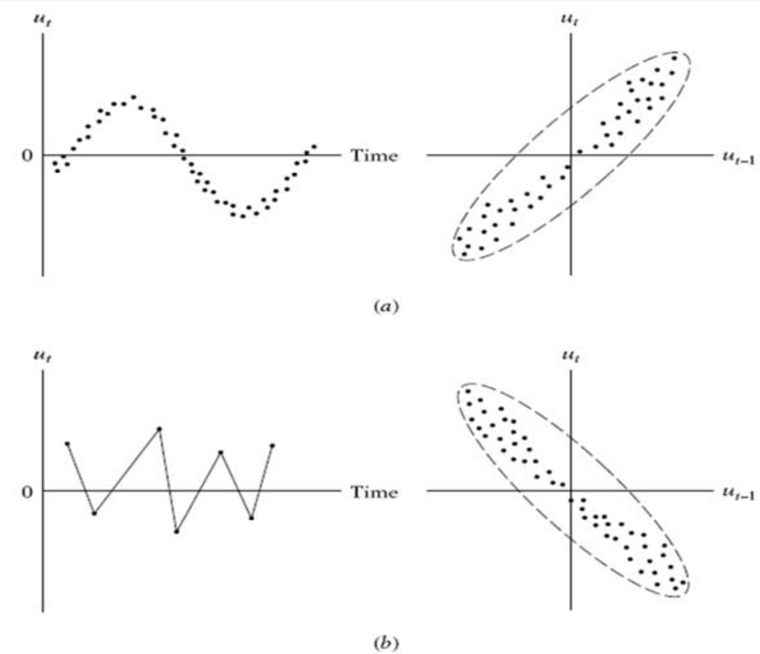

FIGURE 12.3 (*a*) Positive and (*b*) negative autocorrelation.

OLS ESTIMATION IN THE PRESENCE OF AUTOCORRELATION

- **What happens to the OLS estimators and their variances if we introduce autocorrelation in the disturbances by assuming that $E(u_t u_{t+s}) \neq 0$ $(s \neq 0)$ but retain all the other assumptions of the classical model? We revert once again to the two-variable regression model to explain the basic ideas involved, namely,**
- $Y_t = \beta_1 + \beta_2 X_t + u_t$.
- ***To make any headway, we* must assume the *mechanism* that generates u_t*, for* $E(u_t u_{t+s}) \neq 0$ $(s \neq 0)$ *is* too general an assumption to be of any practical use. As a starting point, or first approximation, one can assume that the disturbance, or error, terms are generated by the following mechanism.**
- $u_t = \rho u_{t-1} + \varepsilon_t \qquad -1 < \rho < 1 \qquad (12.2.1)$
- **where *ρ (= rho) is known as the coefficient of autocovariance and where ε_t is the stochastic disturbance term such that it satisfied the standard OLS* assumptions, namely,**
- $E(\varepsilon t) = 0$
- $\text{var}(\varepsilon_t) = \sigma^2\varepsilon \qquad (12.2.2)$
- $\text{cov}(\varepsilon_t, \varepsilon_{t+s}) = 0 \quad s \neq 0$

- **In the engineering literature, an error term with the preceding properties is often called a *white noise error term*. What (12.2.1) postulates is that the value of the disturbance term in period *t is equal to rho times its value in the* previous period plus *a purely random error term*.**
- **The scheme (12.2.1) is known as *Markov first-order autoregressive***
- ***scheme,* or simply a first-order autoregressive scheme, usually denoted as AR(1). *It is first order* because u_t *and its immediate past value are involved; that is, the maximum* lag is 1. If the model were $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$ *, it would be an AR(2), or* second-order, autoregressive scheme, and so on.**
- **In passing, note that rho, the coefficient of autocovariance in (12.2.1), can also be interpreted as the first-order coefficient of autocorrelation, or more accurately, *the coefficient of autocorrelation at lag 1*.**

- **Given the AR(1) scheme, it can be shown that (see Appendix 12A, Section**
- **12A.2)**

$$\text{var}(u_t) = E(u_t^2) = \frac{\sigma_\varepsilon^2}{1-\rho^2} \qquad (12.2.3)$$

$$\text{cov}(u_t, u_{t+s}) = E(u_t u_{t-s}) = \rho^s \frac{\sigma_\varepsilon^2}{1-\rho^2} \qquad (12.2.4)$$

$$\text{cor}(u_t, u_{t+s}) = \rho^s \qquad (12.2.5)$$

- **Since *ρ is a constant between −1 and +1, (12.2.3) shows that under the* AR(1) scheme, the variance of u_t *is still homoscedastic, but* u_t *is correlated* not only with its immediate past value but its values several periods in the past. It is *critical to note that |ρ| < 1, that is, the absolute value of rho is less* than one. If, for example, rho is one, the variances and covariances listed above are not defined.**

- **If |ρ| *< 1, we say that the AR(1) process given in* (12.2.1) *is stationary; that is, the mean, variance, and covariance of* u_t *do not* change over time. If |ρ| *is less than one, then it is clear from (12.2.4) that the* value of the covariance will decline as we go into the distant past.**
- **One reason we use the AR(1) process is not only because of its simplicity compared to higher-order AR schemes, but also because in many applications it has proved to be quite useful. Additionally, a considerable amount of theoretical and empirical work has been done on the AR(1) scheme.**
- **Now return to our two-variable regression model: $Y_t = \beta_1 + \beta_2 X_t + u_t$. *We* know from Chapter 3 that the OLS estimator of the slope coefficient is**

$$\hat{\beta}_2 = \frac{\sum x_t y_t}{\sum x_t^2} \qquad (12.2.6)$$

- **and its variance is given by**

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_t^2} \qquad (12.2.7)$$

- **where the small letters as usual denote deviation from the mean values.**

- **Now under the AR(1) scheme, it can be shown that the variance of this estimator is:**

$$\operatorname{var}(\hat{\beta}_2)_{\text{AR1}} = \frac{\sigma^2}{\sum x_t^2}\left[1 + 2\rho\frac{\sum x_t x_{t-1}}{\sum x_t^2} + 2\rho^2\frac{\sum x_t x_{t-2}}{\sum x_t^2} + \cdots + 2\rho^{n-1}\frac{x_1 x_n}{\sum x_t^2}\right] \quad \textbf{(12.2.8)}$$

- **A comparison of (12.2.8) with (12.2.7) shows the former is equal to the latter times a term that depends on** ***ρ as well as the sample autocorrelations*** **between the values taken by the regressor** ***X at various lags. And in general*** **we cannot foretell whether var (*****βˆ2) is less than or greater than var (βˆ2)AR1*** **[but see Eq. (12.4.1) below]. Of course, if rho is zero, the two formulas will coincide, as they should (why?). Also, if the correlations among the successive values of the regressor are very small, the usual OLS variance of the slope estimator will not be seriously biased. But, as a general principle, the two variances will not be the same.**

- **To give some idea about the difference between the variances given in (12.2.7) and (12.2.8), assume that the regressor** ***X also follows the first-order*** **autoregressive scheme with a coefficient of autocorrelation of** ***r. Then it can*** **be shown that (12.2.8) reduces to:**

$$\operatorname{var}(\hat{\beta}_2)_{\text{AR(1)}} = \frac{\sigma^2}{\sum x_t^2}\left(\frac{1+r\rho}{1-r\rho}\right) = \operatorname{var}(\hat{\beta}_2)_{\text{OLS}}\left(\frac{1+r\rho}{1-r\rho}\right) \quad \textbf{(12.2.9)}$$

- **var (*****βˆ2)AR(1) = σ2 x2 t 1 + rρ*** **1 − rρ** ***= var (βˆ2)OLS 1 + rρ*** **1 − rρ** ***(12.2.9)***
- **If, for example,** ***r = 0.6 and ρ = 0.8, using (12.2.9) we can check that*** **var (*****βˆ2)AR1 = 2.8461 var (βˆ2)OLS. To put it another way, var (βˆ2)OLS = 1 2.8461var (βˆ2)AR1 = 0.3513 var (βˆ2)AR1 . That is, the usual OLS formula [i.e.,*** **(12.2.7)] will underestimate the variance of (*****βˆ2)AR1 by about 65 percent. As*** **you will realize, this answer is specific for the given values of** ***r and ρ. But the*** **point of this exercise is to warn you that a blind application of the usual OLS formulas to compute the variances and standard errors of the OLS estimators could give seriously misleading results.**

- ***What*** **now are the properties of** $\hat{\beta}_2$**?** $\hat{\beta}_2$ ***is still linear and*** **unbiased.** ***Is*** $\hat{\beta}_2$ ***still BLUE? Unfortunately,*** **it is not; in the class of linear unbiased estimators, it does not have minimum variance.**

RELATIONSHIP BETWEEN WAGES AND PRODUCTIVITY IN THE BUSINESS SECTOR OF THE UNITED STATES, 1959–1998

- **Now that we have discussed the consequences of autocorrelation, the obvious question is, How do we detect it and how do we correct for it?**
- **Before we turn to these topics, it is useful to consider a concrete example. Table 12.4 gives data on indexes of real compensation per hour (*****Y) and output per hour*** **(*****X) in the business sector of the U.S. economy for the period 1959–1998, the*** **base of the indexes being 1992 = 100.**
- **First plotting the data on** ***Y and X, we obtain Figure 12.7. Since the relationship*** **between real compensation and labor productivity is expected to be positive, it is not surprising that the two variables are positively related. What is surprising is that the relationship between the two is almost linear, although there is some hint that at higher values of productivity the relationship between the two may be slightly nonlinear.**

TABLE 12.4 INDEXES OF REAL COMPENSATION AND PRODUCTIVITY, UNITED STATES, 1959–1998

Observation	Y	X	Observation	Y	X
1959	58.5	47.2	1979	90.0	79.7
1960	59.9	48.0	1980	89.7	79.8
1961	61.7	49.8	1981	89.8	81.4
1962	63.9	52.1	1982	91.1	81.2
1963	65.3	54.1	1983	91.2	84.0
1964	67.8	54.6	1984	91.5	86.4
1965	69.3	58.6	1985	92.8	88.1
1966	71.8	61.0	1986	95.9	90.7
1967	73.7	62.3	1987	96.3	91.3
1968	76.5	64.5	1988	97.3	92.4
1969	77.6	64.8	1989	95.8	93.3
1970	79.0	66.2	1990	96.4	94.5
1971	80.5	68.8	1991	97.4	95.9
1972	82.9	71.0	1992	100.0	100.0
1973	84.7	73.1	1993	99.9	100.1
1974	83.7	72.2	1994	99.7	101.4
1975	84.5	74.8	1995	99.1	102.2
1976	87.0	77.2	1996	99.6	105.2
1977	88.1	78.4	1997	101.1	107.5
1978	89.7	79.5	1998	105.1	110.5

Notes: X = index of output per hour, business sector (1992 = 100)
Y = index of real compensation per hour, business sector (1992 = 100)
Source: Economic Report of the President, 2000, Table B–47, p. 362.

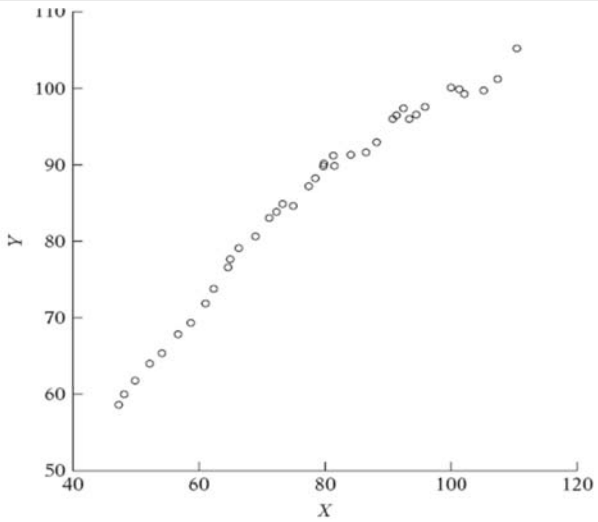

FIGURE 12.7 Index of compensation (Y) and index of productivity (X), United States, 1959–1998.

- **Therefore, we decided to estimate a linear as well as a log–linear model, with the following results:**
- $\hat{Y}_t = 29.5192 + 0.7136X_t$
- **se = (1.9423) (0.0241)**
- $t =$ *(15.1977) (29.6066)* *(12.5.1)*
- $r^2 = 0.9584$ $\quad d = 0.1229$ $\quad \hat{\sigma} = 2.6755$

- **where *d is the Durbin–Watson statistic, which will be discussed shortly.***

- **ln** $Y_t = 1.5239 + 0.6716 \ln X_t$
- **se = (0.0762) (0.0175)**
- t *= (19.9945) (38.2892)* *(12.5.2)*
- $r^2 = 0.9747$ $\quad d = 0.1542$ $\quad \hat{\sigma} = 0.0260$

- **Qualitatively, both the models give similar results. In both cases the estimated coefficients are "highly" significant, as indicated by the high *t values.***
- **In the linear model, if the index of productivity goes up by a unit, on average, the index of compensation goes up by about 0.71 units. In the log–linear model, the slope coefficient being elasticity, we find that if the index of productivity goes up by 1 percent, on average, the index of real compensation goes up by about 0.67 percent.**
- **How reliable are the results given in (12.5.1) and (12.5.2) if there is autocorrelation? As stated previously, if there is autocorrelation, the *estimated standard errors are biased*, as a result of which the *estimated t ratios are unreliable.* We obviously need to find out if our data suffer from autocorrelation. In the following section we discuss several methods of detecting autocorrelation.**

DETECTING AUTOCORRELATION

- **I. Graphical Method**
- **Recall that the assumption of nonautocorrelation of the classical model relates to the population disturbances u_t, *which are not directly observable.* What we have instead are their proxies, the residuals $\hat{u}_t$, *which can be obtained* by the usual OLS procedure. Although the *$\hat{u}_t$ are not the same thing* as u_t ,*17 very often a visual examination of the $\hat{u}$'s gives us some clues about* the likely presence of autocorrelation in the *u's. Actually, a visual examination* of *$\hat{u}_t$ or ($\hat{u}^2_t$)* can provide useful information about autocorrelation, model inadequacy, or specification bias.**
- **There are various ways of examining the residuals. We can simply plot them against time, *the time sequence plot*, as we have done in Figure 12.8, which shows the residuals obtained from the wages–productivity regression (12.5.1). The values of these residuals are given in Table 12.5 along with some other data.**

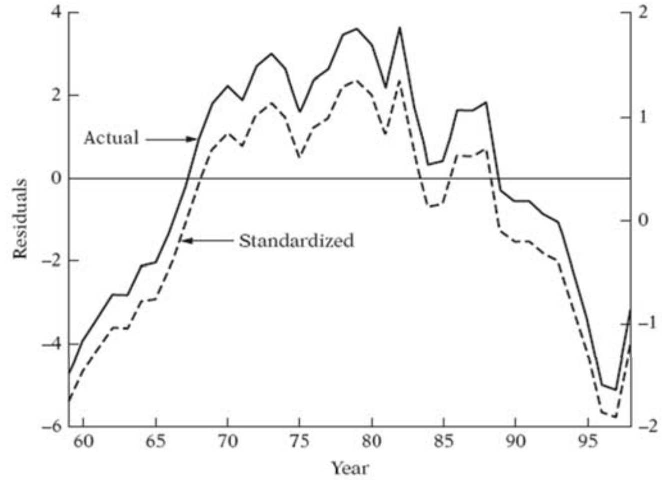

FIGURE 12.8 Residuals and standardized residuals from the wages–productivity regression (12.5.1).

TABLE 12.5 RESIDUALS: ACTUAL, STANDARDIZED, AND LAGGED

Observation	RES1	SRES1	RES1(−1)	Observation	RES1	SRES1	RES1(−1)
1959	−4.703979	−1.758168		1979	3.602089	1.346324	3.444821
1960	−3.874907	−1.448293	−4.703979	1980	3.230723	1.207521	3.602089
1961	−3.359494	−1.255651	−3.874907	1981	2.188868	0.818116	3.230723
1962	−2.800911	−1.046874	−3.359494	1982	3.631600	1.357354	2.188868
1963	−2.828229	−1.057084	−2.800911	1983	1.733354	0.647862	3.631600
1964	−2.112378	−0.789526	−2.828229	1984	0.320571	0.119817	1.733354
1965	−2.039697	−0.762361	−2.112378	1985	0.407350	0.152252	0.320571
1966	−1.252480	−0.468129	−2.039697	1986	1.651836	0.617393	0.407350
1967	−0.280237	−0.104742	−1.252480	1987	1.623640	0.606855	1.651836
1968	0.949713	0.354966	−0.280237	1988	1.838615	0.687204	1.623640
1969	1.835615	0.686083	0.949713	1989	−0.303679	−0.113504	1.838615
1970	2.236492	0.835915	1.835615	1990	−0.560070	−0.209333	−0.303679
1971	1.880977	0.703038	2.236492	1991	−0.559193	−0.209005	−0.560070
1972	2.710926	1.013241	1.880977	1992	−0.885197	−0.330853	−0.559193
1973	3.012241	1.125861	2.710926	1993	−1.056563	−0.394903	−0.885197
1974	2.654535	0.992164	3.012241	1994	−2.184320	−0.816416	−1.056563
1975	1.599020	0.597653	2.654535	1995	−3.355248	−1.254064	−2.184320
1976	2.386238	0.891885	1.599020	1996	−4.996226	−1.867399	−3.355248
1977	2.629847	0.982936	2.386238	1997	−5.137643	−1.920255	−4.996226
1978	3.444821	1.287543	2.629847	1998	−3.278621	−1.225424	−5.137643

Notes: RES 1 = residuals from regression (12.5.1).
SRES 1 = standardized residuals = RES1/2.6755.
RES(−1) = residuals lagged one period.

- **To see this differently, we can plot $\hat{u}_t$ *against $\hat{u}_{t-1}$, that is, plot the residuals* at time *t against their value at time (t − 1), a kind of empirical test of the* AR(1) scheme. If the residuals are nonrandom, we should obtain pictures similar to those shown in Figure 12.3. This plot for our wages–productivity regression is as shown in Figure 12.9; the underlying data are given in**

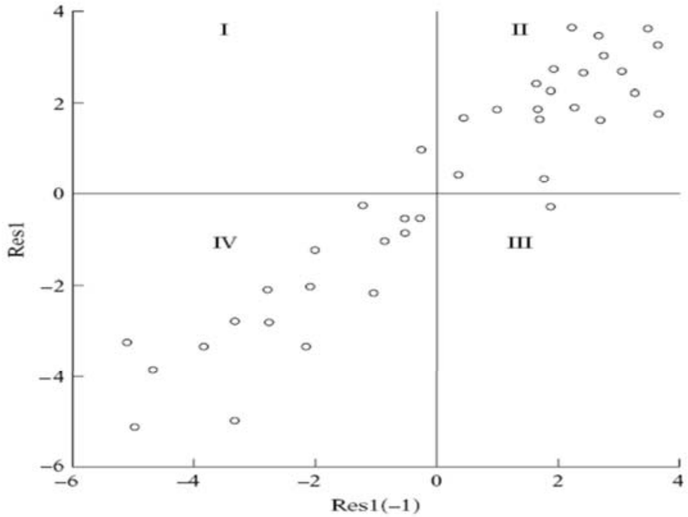

FIGURE 12.9 Current residuals versus lagged residuals.

- **II. The Runs Test**
- **If we carefully examine Figure 12.8, we notice a peculiar feature: Initially, we have several residuals that are negative, then there is a series of positive residuals, and then there are several residuals that are negative. If these residuals were purely random, could we observe such a pattern? Intuitively, it seems unlikely. This intuition can be checked by the so-called runs test, sometimes also know as the Geary test, a nonparametric test.**
- **To explain the runs test, let us simply note down the signs (+ or −) of the residuals obtained from the wages–productivity regression, which are given in the first column of Table 12.5.**

- **(−−−−−−−−−)(+++++++++++++++++++++)(−−−−−−−−−−) (12.6.1)**

- **Thus there are 9 negative residuals, followed by 21 positive residuals, followed by 10 negative residuals, for a total of 40 observations.**
- **We now define a run as an uninterrupted sequence of one symbol or attribute, such as + or −. We further define the length of a run as the number of elements in it. In the sequence shown in (12.6.1), there are 3 runs: a run of 9 minuses (i.e., of length 9), a run of 21 pluses (i.e., of length 21) and a run of 10 minuses (i.e., of length 10). For a better visual effect, we have presented the various runs in parentheses.**
- **By examining how runs behave in a strictly random sequence of observations, one can derive a test of randomness of runs. We ask this question: Are the 3 runs observed in our illustrative example consisting of 40 observations too many or too few compared with the number of runs expected in a strictly random sequence of 40 observations? If there are too many runs, it would mean that in our example the residuals change sign frequently, thus indicating negative serial correlation (cf. Figure 12.3*b*). *Similarly, if there* are too few runs, they may suggest positive autocorrelation, as in Figure 12.3*a. A priori, then, Figure 12.8 would indicate positive correlation in* the residuals.**

- **Now let**
- ***N = total number of observations = N1 + N2***
- ***N1 = number of + symbols (i.e., + residuals)***
- ***N2 = number of − symbols (i.e., − residuals)***
- ***R = number of runs***

$$\text{Mean:}\quad E(R) = \frac{2N_1N_2}{N} + 1$$
$$\text{Variance:}\quad \sigma_R^2 = \frac{2N_1N_2(2N_1N_2 - N)}{(N)^2(N-1)} \qquad \textbf{(12.6.2)}$$

- ***Note: N = N1 + N2.***
- **If the null hypothesis of randomness is sustainable, following the properties of the normal distribution, we should expect that**
  Prob [*E(R) − 1.96σR ≤ R ≤ E(R) + 1.96σR] = 0.95 (12.6.3)*

Decision Rule. Do not reject the null hypothesis of randomness with 95% confidence if R, the number of runs, lies in the preceding confidence interval; reject the null hypothesis if the estimated R lies outside these limits. (*Note:* You can choose any level of confidence you want.)

- **Using the formulas given in (12.6.2), we obtain**

$$\begin{aligned} E(R) &= 10.975 \\ \sigma_R^2 &= 9.6936 \qquad\qquad (12.6.4) \\ \sigma_R &= 3.1134 \end{aligned}$$

- **The 95% confidence interval for *R in our example is thus:***

$[10.975 \pm 1.96(3.1134)] = (4.8728, 17.0722)$

- **Durbin–Watson *d Test***
- **The most celebrated test for detecting serial correlation is that developed**
- **by statisticians Durbin and Watson. It is popularly known as the Durbin–**
- **Watson *d statistic, which is defined as***

$$d = \frac{\sum_{t=2}^{t=n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{t=n} \hat{u}_t^2} \qquad (12.6.5)$$

- **it is important to note the assumptions underlying the *d statistic.***
- **1. The regression model includes the intercept term. If it is not present, as in the case of the regression through the origin, it is essential to rerun the regression including the intercept term to obtain the RSS.**
- **2. The explanatory variables, the *X's, are nonstochastic, or fixed in repeated* sampling.**

- **3. The disturbances *ut are generated by the first-order autoregressive* scheme: $u_t = \rho u_{t-1} + \varepsilon_t$. *Therefore, it cannot be used to detect higher-order* autoregressive schemes.**
- **4. The error term u_t *is assumed to be normally distributed.***
- **5. The regression model does not include the lagged value(s) of the dependent variable as one of the explanatory variables. Thus, the test is inapplicable in models of the following type:**
- $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + \beta_k X_{kt} + \gamma Y_{t-1} + u_t$ *(12.6.6)*
- **where *Yt−1 is the one period lagged value of Y.***
- **6. There are no missing observations in the data. Thus, in our wages–productivity regression for the period 1959–1998, if observations for, say, 1978 and 1982 were missing for some reason, the *d statistic makes no* allowance for such missing observations**

- $d \approx 2(1-\hat{\rho})$ *(12.6.10)*
- **But since $-1 \le \rho \le 1$, *(12.6.10) implies that***

$0 \le d \le 4$ *(12.6.11)*
- **These are the bounds of *d; any estimated d value must lie within these***
- **limits.**

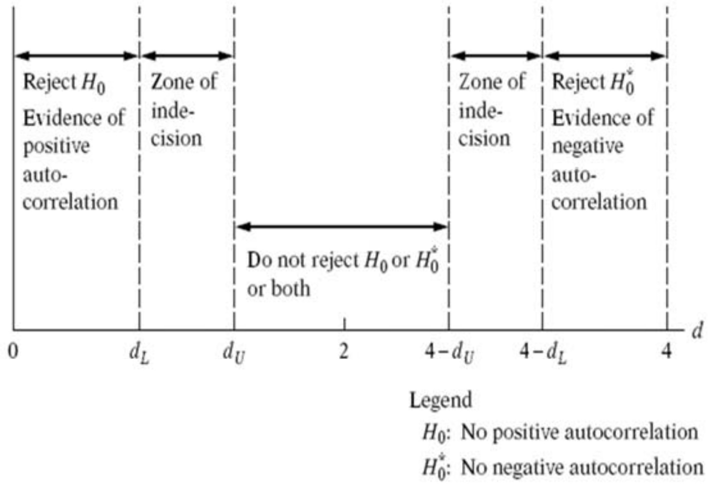

FIGURE 12.10 Durbin–Watson *d* statistic.

- **The mechanics of the Durbin–Watson test are as follows, assuming that the assumptions underlying the test are fulfilled:**
- **1. Run the OLS regression and obtain the residuals.**
- **2. Compute *d from (12.6.5). (Most computer programs now do this* routinely.)**
- **3. For the given sample size and given number of explanatory variables, find out the critical *dL and dU values.***
- **4. Now follow the decision rules given in Table 12.6. For ease of reference, these decision rules are also depicted in Figure 12.10.**

- **To illustrate the mechanics, let us return to our wages–productivity regression. From the data given in Table 12.5 the estimated *d value can be* shown to be 0.1229, suggesting that there is positive serial correlation in the residuals. From the Durbin–Watson tables, we find that for 40 observations and one explanatory variable, *dL = 1.44 and dU = 1.54 at the 5 percent level.* Since the computed *d of 0.1229 lies below dL, we cannot reject the hypothesis* that there is positive serial correlations in the residuals.**