Chapter 12

Regression with Time-Series Data: Nonstationary Variables

Learning Objectives

Based on the material in this chapter, you should be able to

1. Explain the differences between stationary and nonstationary time-series processes.
2. Describe the general behavior of an autoregressive process and a random walk process.
3. Explain why we need "unit root" tests, and state implications of the null and alternative hypotheses.
4. Explain what is meant by the statement that a series is "integrated of order one" or I(1).
5. Perform Dickey–Fuller and augmented Dickey–Fuller tests for stationarity.
6. Explain the meaning of a "spurious regression".
7. Explain the concept of cointegration and test whether two series are cointegrated.
8. Explain how to choose an appropriate model for regression analysis with time-series data.

Keywords

autoregressive process	order of integration	stochastic process
cointegration	random walk process	stochastic trend
Dickey–Fuller tests	random walk with drift	tau statistic
difference stationary	spurious regressions	trend stationary
mean reversion	stationary	unit root tests
nonstationary		

In 2003 the Nobel Prize in Economics[1] was awarded jointly to two distinguished econometricians: Professor Robert F. Engle "for methods of analyzing economic time series with time-varying volatility (ARCH)" and Professor Clive W. J. Granger "for

[1] For more details, see http://nobelprize.org/nobel_prizes/economics/.

methods of analyzing economic time series with common trends (cointegration)." The aim of this and the following two chapters is to discuss the background that prompted these contributions, and to show how the proposed methods have revolutionized the way we conduct econometrics with time-series data.

The analysis of time-series data is of vital interest to many groups, such as macroeconomists studying the behavior of national and international economies, finance economists analyzing the stock market, and agricultural economists predicting supplies and demands for agricultural products. For example, if we are interested in forecasting the growth of gross domestic product or inflation, we look at various indicators of economic performance and consider their behavior over recent years. Alternatively, if we are interested in a particular business, we analyze the history of the industry in an attempt to predict potential sales. In each of these cases, we are analyzing time-series data.

We have already worked with time-series data in Chapter 9 and have discovered how regression models for these data often have special characteristics designed to capture their dynamic nature. We saw how including lagged values of the dependent variable or explanatory variables as regressors, or considering lags in the errors, can be used to model dynamic relationships. We also showed how autoregressive models can be used in forecasting. However, an important assumption maintained throughout Chapter 9 was that the variables have a property called stationarity. It is time now to learn the difference between stationary and nonstationary variables. Many economic variables are nonstationary and, as you will learn, the consequences of nonstationary variables for regression modeling are profound.

The aim of this chapter is to describe how to estimate regression models involving nonstationary variables. The first step in this direction is to examine the time-series concepts of **stationarity** (and **nonstationarity**) and how we distinguish between them. **Cointegration** is another important related concept that has a bearing on our choice of a regression model. The seminal contributions of the Nobel laureates show that the econometric consequences of nonstationarity can be quite severe, and offer methods to overcome them.

12.1 Stationary and Nonstationary Variables

Plots of the time series of some important economic variables for the U.S. economy are displayed in Figure 12.1. The data for these figures can be found in the file *usa.dat*. The figures on the left-hand side are the real gross domestic product (a measure of aggregate economic production), the annual inflation rate (a measure of changes in the aggregate price level), the federal funds rate (the interest rate on overnight loans between banks), and the three-year bond rate (interest rate on a financial asset to be held for three years). Observe how the GDP variable displays upward trending behavior, while the inflation rate appears to "wander up and down" with no discernable pattern or trend. Similarly, both the federal funds rate and the bond rate show "wandering up and down" behavior. The figures on the right-hand side of Figure 12.1 are the changes of the corresponding variables on the left-hand side.

The change in a variable is an important concept that is used repeatedly in this chapter; it is worth dwelling on its definition. The change in a variable y_t, also known as its first difference, is given by $\Delta y_t = y_t - y_{t-1}$. Thus Δy_t is the change in the value of the variable y from period $t-1$ to period t.

The time series of the changes on the right-hand side of Figure 12.1 display behavior that can be described as irregular ups and downs, or fluctuations. Note that while changes in the inflation rate and the two interest rates appear to fluctuate around a constant value, the

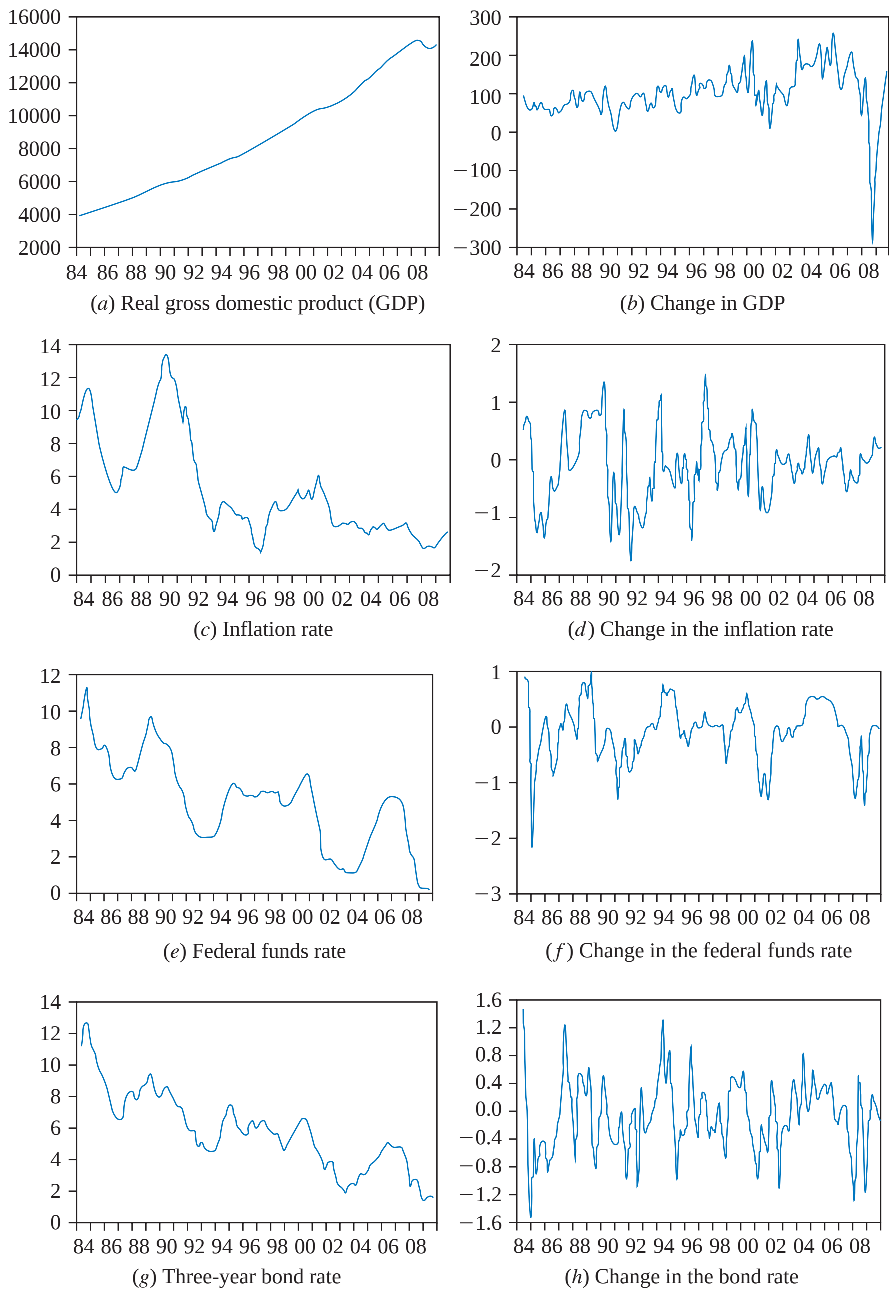

FIGURE 12.1 U.S. economic time series.

changes in the GDP variable appear to fluctuate around an upward trend, until the financial crisis. The first question we address in this chapter is: Which data series represent stationary variables and which are observations on nonstationary variables?

Formally, a time series y_t is stationary if its mean and variance are constant over time, and if the covariance between two values from the series depends only on the length of time separating the two values, and not on the actual times at which the variables are observed. That is, the time series y_t is stationary if for all values, and every time period, it is true that

Table 12.1 Sample Means of Time Series Shown in Figure 12.1

Variable	Sample periods 1984:2 to 1996:4	Sample periods 1997:1 to 2009:4
Real GDP (a)	5813.0	11458.2
Inflation rate (c)	6.9	3.2
Federal funds rate (e)	6.4	3.5
Bond rate (g)	7.3	4.0
Change in GDP (b)	82.7	120.3
Change in the inflation rate (d)	−0.16	0.02
Change in the federal funds rate (f)	−0.09	−0.10
Change in the bond rate (h)	−0.10	−0.09

$$E(y_t) = \mu \quad \text{(constant mean)} \tag{12.1a}$$

$$\operatorname{var}(y_t) = \sigma^2 \quad \text{(constant variance)} \tag{12.1b}$$

$$\operatorname{cov}(y_t, y_{t+s}) = \operatorname{cov}(y_t, y_{t-s}) = \gamma_s \quad \text{(covariance depends on } s, \text{ not } t) \tag{12.1c}$$

The first condition, that of a constant mean, is the feature that has received the most attention. To appreciate this condition for stationarity, look at the plots shown in Figure 12.1 and their sample means shown in Table 12.1. The sample means for the changes in the two interest rates are similar across different sample periods, whereas the sample means for the variables in the original levels, as well as the changes in GDP and inflation, differ across sample periods. Thus, while the federal funds rate, and the bond rate display characteristics of nonstationarity, their changes display characteristics of stationarity. For inflation and GDP, both their levels and their changes display characteristics of nonstationarity. Nonstationary series with nonconstant means are often described as *not* having the property of **mean reversion**. That is, stationary series have the property of mean reversion.

Looking at the sample means of time-series variables is a convenient indicator as to whether a series is stationary or nonstationary, but this does not constitute a hypothesis test. A formal test is described in Section 12.3. However, before we introduce the test, it is useful to revisit the first-order autoregressive model that was introduced in Chapter 9.

12.1.1 The First-Order Autoregressive Model

Let y_t be an economic variable that we observe over time. In line with most economic variables, we assume that y_t is random, since we cannot perfectly predict it. We never know the values of random variables until they are observed. The econometric model generating a time-series variable y_t is called a **stochastic** or **random process**. A sample of observed y_t values is called a particular **realization** of the stochastic process. It is one of many possible paths that the stochastic process could have taken. Univariate time-series models are examples of stochastic processes where a single variable y is related to past values of itself and current and past error terms. In contrast to regression modeling, univariate time-series models do not contain any explanatory variables (no x's).

The autoregressive model of order one, the AR(1) model, is a useful univariate time-series model for explaining the difference between stationary and nonstationary series. It is given by

$$y_t = \rho y_{t-1} + v_t, \quad |\rho| < 1 \tag{12.2a}$$

where the errors v_t are independent, with zero mean and constant variance σ_v^2, and may be normally distributed. In the context of time-series models, the errors are sometimes known as "shocks" or "innovations." As we will see, the assumption $|\rho| < 1$ implies that y_t is stationary. The AR(1) process shows that each realization of the random variable y_t contains a proportion ρ of last period's value y_{t-1} plus an error v_t drawn from a distribution with mean zero and variance σ_v^2. Since we are concerned with only one lag, the model is described as an autoregressive model of order one. In general an AR(p) model includes lags of the variable y_t up to y_{t-p}. An example of an AR(1) time series with $\rho = 0.7$, and independent $N(0,1)$ random errors is shown in Figure 12.2a. Note that the data have been artificially generated. Observe how the time series fluctuates around zero and has no trend-like behavior, a characteristic of stationary series.

The value "zero" is the constant mean of the series, and it can be determined by doing some algebra known as recursive substitution.[2] Consider the value of y at time $t = 1$, then its value at time $t = 2$ and so on. These values are

$$\begin{aligned}
y_1 &= \rho y_0 + v_1 \\
y_2 &= \rho y_1 + v_2 = \rho(\rho y_0 + v_1) + v_2 = \rho^2 y_0 + \rho v_1 + v_2 \\
&\vdots \\
y_t &= v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \cdots + \rho^t y_0
\end{aligned}$$

The mean of y_t is

$$E(y_t) = E(v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \cdots) = 0$$

since the error v_t has zero mean and the value of $\rho^t y_0$ is negligible for a large t. The variance can be shown to be a constant $\sigma_v^2/(1-\rho^2)$ while the covariance between two errors s periods apart γ_s can be shown to be $\sigma_v^2 \rho^s/(1-\rho^2)$. Thus, the AR(1) model in (12.2a) is a classic example of a stationary process with a zero mean.

Real-world data rarely have a zero mean. We can introduce a nonzero mean μ by replacing y_t in (12.2a) with $(y_t - \mu)$ as follows:

$$(y_t - \mu) = \rho(y_{t-1} - \mu) + v_t$$

which can then be rearranged as

$$y_t = \alpha + \rho y_{t-1} + v_t, \quad |\rho| < 1 \tag{12.2b}$$

[2] An alternative to recursive substitution when the variable is stationary is to use the lag operator algebra discussed in Chapter 9.8.

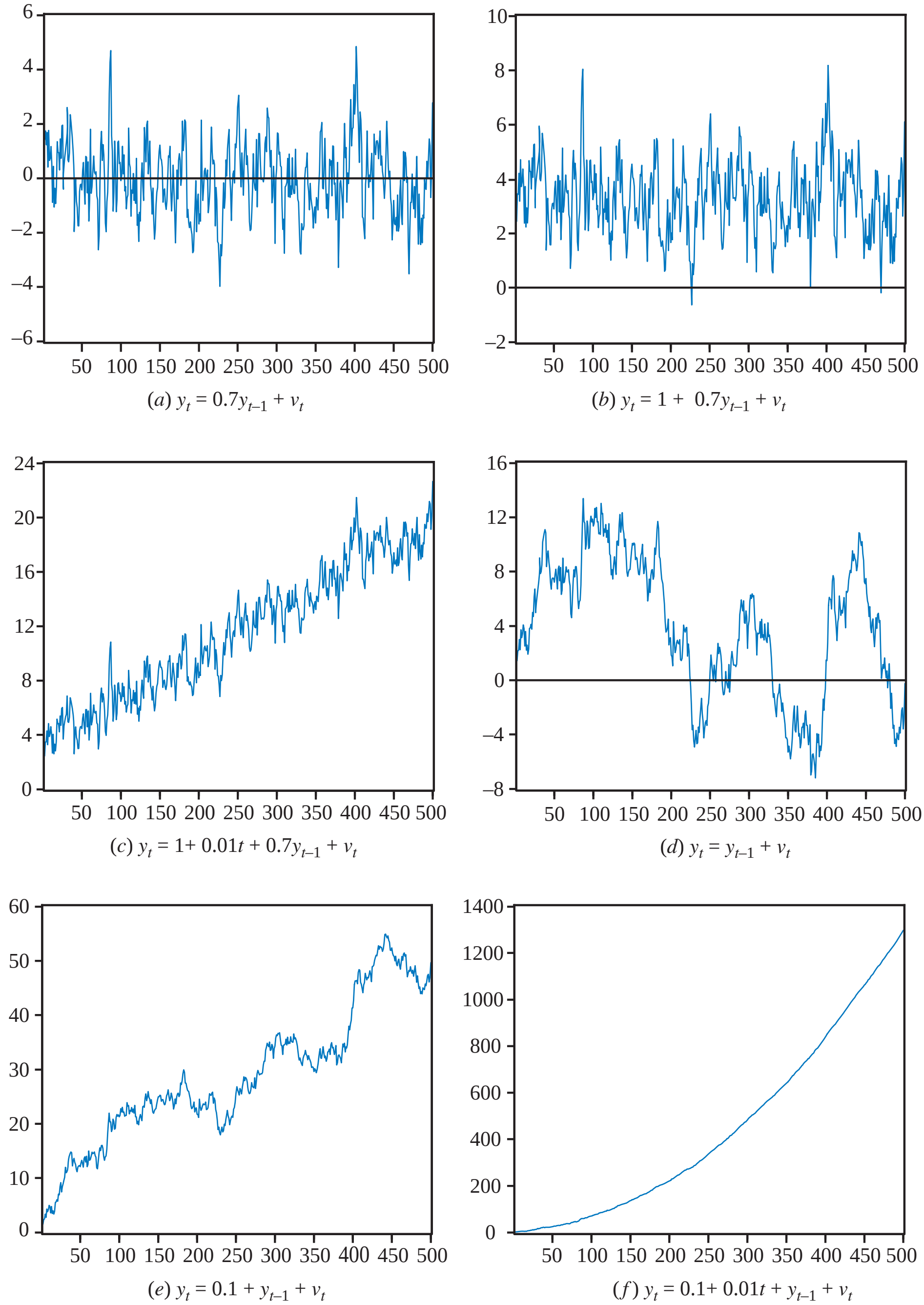

FIGURE 12.2 Time-series models.

where $\alpha = \mu(1-\rho)$. That is, we can accommodate a nonzero mean in y_t by either working with the "de-meaned" variable $(y_t - \mu)$ or introducing the intercept term α in the autoregressive process of y_t as in (12.2b). Corresponding to these two ways, we describe the "de-meaned" variable $(y_t - \mu)$ as being stationary around zero, or the variable y_t as stationary around its mean value $\mu = \alpha/(1-\rho)$.

An example of a time series that follows this model, with $\alpha = 1$, $\rho = 0.7$ is shown in Figure 12.2(b). We have used the same values of the error v_t as in Figure 12.2(a), so the figure shows the added influence of the constant term. Note that the series now fluctuates around a nonzero value. This nonzero value is the constant mean of the series

$$E(y_t) = \mu = \alpha/(1-\rho) = 1/(1-0.7) = 3.33$$

Another extension to (12.2a) is to consider an AR(1) model fluctuating around a linear trend $(\mu + \delta t)$. As we have seen in Figure 12.1, some real-world data appear to exhibit a trend. In this case, we let the "de-trended" series $(y_t - \mu - \delta t)$ behave like an autoregressive model

$$(y_t - \mu - \delta t) = \rho(y_{t-1} - \mu - \delta(t-1)) + v_t, \quad |\rho| < 1$$

which can be rearranged as

$$y_t = \alpha + \rho y_{t-1} + \lambda t + v_t \tag{12.2c}$$

where $\alpha = (\mu(1-\rho) + \rho\delta)$ and $\lambda = \delta(1-\rho)$. An example of a time series that can be described by this model with $\rho = 0.7$, $\alpha = 1$, and $\delta = 0.01$ is shown in Figure 12.2(c). The de-trended series $(y_t - \mu - \delta t)$ also has a constant variance and covariances that depend only on the time separating observations, not the time at which they are observed. In other words, the "de-trended" series is stationary. An astute reader may have noted that the mean of y_t, $E(y_t) = \mu + \delta t$ depends on t, which implies that y_t is nonstationary. While this observation is correct, when $|\rho| < 1$, y_t is more usually described as stationary around the deterministic trend line $\mu + \delta t$. This is discussed further in Section 12.5.2.

12.1.2 Random Walk Models

Consider the special case of $\rho = 1$ in (12.2a):

$$y_t = y_{t-1} + v_t \tag{12.3a}$$

This model is known as the random walk model. Equation (12.3a) shows that each realization of the random variable y_t contains last period's value y_{t-1} plus an error v_t. An example of a time series that can be described by this model is shown in Figure 12.2(d). These time series are called **random walks** because they appear to wander slowly upward or downward with no real pattern; the values of sample means calculated from subsamples of observations will be dependent on the sample period. This is a characteristic of nonstationary series.

We can understand the "wandering" behavior of random walk models by doing some recursive substitution.

$$\begin{aligned}
y_1 &= y_0 + v_1 \\
y_2 &= y_1 + v_2 = (y_0 + v_1) + v_2 = y_0 + \sum_{s=1}^{2} v_s \\
&\vdots \\
y_t &= y_{t-1} + v_t = y_0 + \sum_{s=1}^{t} v_s
\end{aligned}$$

The random walk model contains an initial value y_0 (often set to zero because it is so far in the past that its contribution to y_t is negligible) plus a component that is the sum of the past stochastic terms $\sum_{s=1}^{t} v_s$. This latter component is often called the **stochastic trend**. This term arises because a stochastic component v_t is added for each time t, and because it causes the time series to trend in unpredictable directions. If the variable y_t is subjected to a sequence of positive shocks, $v_t > 0$, followed by a sequence of negative shocks, $v_t < 0$, it will have the appearance of wandering upward, then downward.

We have used the fact that y_t is a sum of errors to explain graphically the nonstationary nature of the random walk. We can also use it to show algebraically that the conditions for stationarity do not hold. Recognizing that the v_t are independent, taking the expectation and the variance of y_t yields, for a fixed initial value y_0,

$$E(y_t) = y_0 + E(v_1 + v_2 + \cdots + v_t) = y_0$$
$$\text{var}(y_t) = \text{var}(v_1 + v_2 + \cdots + v_t) = t\sigma_v^2$$

The random walk has a mean equal to its initial value and a variance that increases over time, eventually becoming infinite. Although the mean is constant, the increasing variance implies that the series may not return to its mean, and so sample means taken for different periods are not the same.

Another nonstationary model is obtained by adding a constant term to (12.3a):

$$y_t = \alpha + y_{t-1} + v_t \qquad (12.3b)$$

This model is known as the **random walk with drift**. Equation (12.3b) shows that each realization of the random variable y_t contains an intercept (the drift component α) plus last period's value y_{t-1} plus the error v_t. An example of a time series that can be described by this model (with $\alpha = 0.1$) is shown in Figure 12.2(e). Notice how the time-series data appear to be ''wandering'' as well as ''trending'' upward. In general, random walk with drift models show definite trends either upward (when the drift α is positive) or downward (when the drift α is negative).

Again, we can get a better understanding of this behavior by applying recursive substitution:

$$\begin{aligned}
y_1 &= \alpha + y_0 + v_1 \\
y_2 &= \alpha + y_1 + v_2 = \alpha + (\alpha + y_0 + v_1) + v_2 = 2\alpha + y_0 + \sum_{s=1}^{2} v_s \\
&\vdots \\
y_t &= \alpha + y_{t-1} + v_t = t\alpha + y_0 + \sum_{s=1}^{t} v_s
\end{aligned}$$

The value of y at time t is made up of an initial value y_0, the stochastic trend component $(\sum_{s=1}^{t} v_s)$, and now a **deterministic trend** component $t\alpha$. It is called a deterministic trend because a fixed value α is added for each time t. The variable y wanders up and down as well as increases by a fixed amount at each time t. The mean and variance of y_t are

$$E(y_t) = t\alpha + y_0 + E(v_1 + v_2 + v_3 + \cdots + v_t) = t\alpha + y_0$$
$$\text{var}(y_t) = \text{var}(v_1 + v_2 + v_3 + \cdots + v_t) = t\sigma_v^2$$

In this case both the constant mean and constant variance conditions for stationarity are violated.

We can extend the random walk model even further by adding a time trend:

$$y_t = \alpha + \delta t + y_{t-1} + v_t \tag{12.3c}$$

An example of a time series that can be described by this model (with $\alpha = 0.1$; $\delta = 0.01$) is shown in Figure 12.2(f). Note how the addition of a time-trend variable t strengthens the trend behavior. We can see the amplification using the same algebraic manipulation as before:

$$\begin{aligned}
y_1 &= \alpha + \delta + y_0 + v_1 \\
y_2 &= \alpha + \delta 2 + y_1 + v_2 = \alpha + 2\delta + (\alpha + \delta + y_0 + v_1) + v_2 = 2\alpha + 3\delta + y_0 + \sum_{s=1}^{2} v_s \\
&\vdots \\
y_t &= \alpha + \delta t + y_{t-1} + v_t = t\alpha + \left(\frac{t(t+1)}{2}\right)\delta + y_0 + \sum_{s=1}^{t} v_s
\end{aligned}$$

where we have used the formula for a sum of an arithmetic progression,

$$1 + 2 + 3 + \cdots + t = t(t+1)/2$$

The additional term has the effect of strengthening the trend behavior.

To recap, we have considered the autoregressive class of models and have shown that they display properties of stationarity when $|\rho| < 1$. We have also discussed the random walk class of models when $\rho = 1$. We showed that random walk models display properties of nonstationarity. Now, go back and compare the real-world data in Figure 12.1 with those in Figure 12.2. Ask yourself what models might have generated the different data series in Figure 12.1. In the next few sections we shall consider how to test which series in Figure 12.1 exhibit properties associated with stationarity, as well as which series exhibit properties associated with nonstationarity.

12.2 Spurious Regressions

The main reason why it is important to know whether a time series is stationary or nonstationary before one embarks on a regression analysis is that there is a danger of obtaining apparently significant regression results from unrelated data when nonstationary series are used in regression analysis. Such regressions are said to be **spurious**.

To illustrate the problem, let us take two independent random walks:

$$rw_1:\; y_t = y_{t-1} + v_{1t}$$

$$rw_2:\; x_t = x_{t-1} + v_{2t}$$

where v_{1t} and v_{2t} are independent $N(0,1)$ random errors. Two such series are shown in Figure 12.3(a)—the data are in the file *spurious dat*. These series were generated independently and, in truth, have no relation to one another, yet when we plot them, as we have done in Figure 12.3(b), we see a positive relationship between them. If we estimate a simple regression of series one (rw_1) on series two (rw_2), we obtain the following results:

$$\begin{aligned}
\widehat{rw_{1t}} &= 17.818 + 0.842\, rw_{2t}, \quad R^2 = 0.70 \\
(t) &\qquad\qquad (40.837)
\end{aligned}$$

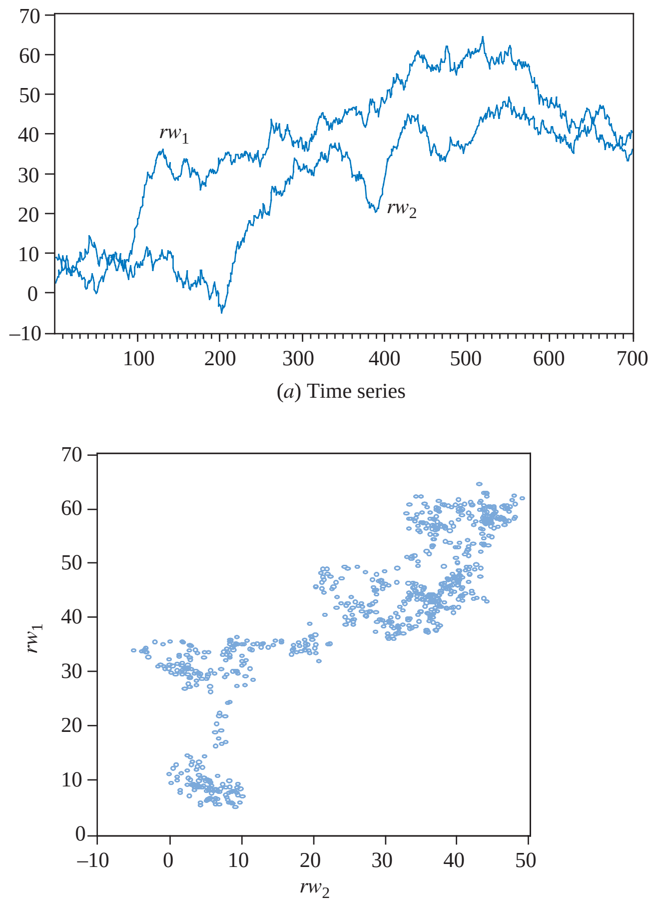

FIGURE 12.3 Time series and scatter plot of two random walk variables.

This result suggests that the simple regression model fits the data well ($R^2 = 0.70$), and that the estimated slope is significantly different from zero. In fact, the *t*-statistic is huge! These results are, however, completely meaningless, or spurious. The apparent significance of the relationship is false. It results from the fact that we have related one series with a stochastic trend to another series with another stochastic trend. In fact, these series have nothing in common, nor are they causally related in any way. Similar and more dramatic results are obtained when random walk with drift series are used in regressions. Typically the residuals from such regressions will be highly correlated. For this example, the *LM* test value to test for first-order autocorrelation (*p*-value in parenthesis) is 682.958 (0.000); a sure sign that there is a problem with the regression.

In other words, when nonstationary time series are used in a regression model, the results may spuriously indicate a significant relationship when there is none. In these cases the least squares estimator and least squares predictor do not have their usual properties, and *t*-statistics are not reliable. Since many macroeconomic time series are nonstationary, it is particularly important to take care when estimating regressions with macroeconomic variables.

How then can we test whether a series is stationary or nonstationary, and how do we conduct regression analysis with nonstationary data? The former is discussed in Section 12.3, while the latter is considered in Section 12.4.

12.3 Unit Root Tests for Stationarity

There are many tests for determining whether a series is stationary or nonstationary. The most popular one, and the one that we discuss, is the Dickey–Fuller test. As noted in our discussion of the autoregressive and random walk models, stochastic processes can include or exclude a constant term and can include or exclude a time trend. There are three variations of the Dickey–Fuller test designed to take account of the role of the constant term and the trend. We begin by describing the test equations and hypotheses for these three cases and then outline the testing procedure.

12.3.1 Dickey–Fuller Test 1 (No Constant and No Trend)

This test is based on the discussion in Section 12.1 where we note that the AR(1) process $y_t = \rho y_{t-1} + v_t$ is stationary when $|\rho| < 1$, but, when $\rho = 1$, it becomes the nonstationary random walk process $y_t = y_{t-1} + v_t$. Hence, one way to test for stationarity is to examine the value of ρ. In other words, we test whether ρ is equal to one or significantly less than one. Tests for this purpose are known as **unit root tests for stationarity**.

To formalize this procedure a little more, consider again the AR(1) model:

$$y_t = \rho y_{t-1} + v_t \tag{12.4}$$

where the v_t are independent random errors with zero mean and constant variance σ_v^2. We can test for nonstationarity by testing the null hypothesis that $\rho = 1$ against the alternative that $|\rho| < 1$, or simply $\rho < 1$. This one-sided (left tail) test is put into a more convenient form by subtracting y_{t-1} from both sides of (12.4) to obtain

$$\begin{aligned} y_t - y_{t-1} &= \rho y_{t-1} - y_{t-1} + v_t \\ \Delta y_t &= (\rho - 1) y_{t-1} + v_t \\ &= \gamma\, y_{t-1} + v_t \end{aligned} \tag{12.5a}$$

where $\gamma = \rho - 1$ and $\Delta y_t = y_t - y_{t-1}$. Then, the hypotheses can be written in terms of either ρ or γ:

$$H_0 : \rho = 1 \Leftrightarrow H_0 : \gamma = 0$$

$$H_1 : \rho < 1 \Leftrightarrow H_1 : \gamma < 0$$

Note that the null hypothesis is that the series is nonstationary. In other words, if we do not reject the null, we conclude that it is a nonstationary process; if we reject the null hypothesis that $\gamma = 0$, then we conclude that the series is stationary.

12.3.2 Dickey–Fuller Test 2 (With Constant but No Trend)

The second Dickey–Fuller test includes a constant term in the test equation:

$$\Delta y_t = \alpha + \gamma y_{t-1} + v_t \tag{12.5b}$$

The null and alternative hypotheses are the same as before. In this case, if we do not reject the null hypothesis that $\gamma = 0$ (or $\rho = 1$), we conclude that the series is nonstationary. If we reject the null hypothesis that $\gamma = 0$, we conclude that the series is stationary.

12.3.3 Dickey–Fuller Test 3 (With Constant and With Trend)

The third Dickey–Fuller test includes a constant and a trend in the test equation:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \lambda t + v_t \tag{12.5c}$$

As before, the null and alternative hypotheses are $H_0:\gamma = 0$ and $H_1:\gamma < 0$. If we do not reject the null hypothesis that $\gamma = 0\ (\rho = 1)$, we conclude that the series is nonstationary. If we reject the null hypothesis that $\gamma = 0$, we conclude that the series is stationary.

12.3.4 The Dickey–Fuller Critical Values

To test the hypothesis in all three cases, we simply estimate the test equation by least squares and examine the t-statistic for the hypothesis that $\gamma = 0$. Unfortunately this t-statistic no longer has the t-distribution that we have used previously to test zero null hypotheses for regression coefficients. A problem arises because when the null hypothesis is true, y_t is nonstationary and has a variance that increases as the sample size increases. This increasing variance alters the distribution of the usual t-statistic when H_0 is true. To recognize this fact, the statistic is often called a τ ***(tau)*** **statistic**, and its value must be compared to specially generated critical values. Note that critical values are generated for the three different tests because, as we have seen in Section 12.1, the addition of the constant term and the time-trend term changes the behavior of the time series.

Originally these critical values were tabulated by the statisticians Professor David Dickey and Professor Wayne Fuller. The values have since been refined, but in deference to the seminal work, unit root tests using these critical values have become known as **Dickey–Fuller tests**. Table 12.2 contains the critical values for the *tau* (τ) statistic for the three cases; they are valid in large samples for a one-tail test.

Note that the Dickey–Fuller critical values are more negative than the standard critical values (shown in the last row). This implies that the τ-statistic must take larger (negative) values than usual for the null hypothesis of nonstationarity $\gamma = 0$ to be rejected in favor of the alternative of stationarity $\gamma < 0$. Specifically, to carry out this one-tail test of significance, if τ_c is the critical value obtained from Table 12.2, we reject the null hypothesis of nonstationarity if $\tau \le \tau_c$. If $\tau > \tau_c$ then we do not reject the null hypothesis that the series y_t is nonstationary. Expressed in a casual way, but one that avoids the proliferation of "double negatives," $\tau \le \tau_c$ suggests that the series is stationary while $\tau > \tau_c$ suggests nonstationarity.

An important extension of the Dickey–Fuller test allows for the possibility that the error term is autocorrelated. Such autocorrelation is likely to occur if our earlier models did not have sufficient lag terms to capture the full dynamic nature of the process. Using the model with an intercept as an example, the extended test equation is

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{s=1}^{m} a_s \Delta y_{t-s} + v_t \tag{12.6}$$

where $\Delta y_{t-1} = (y_{t-1} - y_{t-2})$, $\Delta y_{t-2} = (y_{t-2} - y_{t-3}), \ldots$. We add as many lagged first difference terms as we need to ensure that the residuals are not autocorrelated. As we discovered in Section 9.6, including lags of the dependent variable can be used to eliminate autocorrelation in the errors. The number of lagged terms can be determined by examining the autocorrelation function (ACF) of the residuals v_t, or the significance of the estimated lag coefficients a_s. The unit root tests based on (12.6) and its variants (intercept excluded or trend included) are referred to as **augmented Dickey–Fuller tests**. The hypotheses for stationarity

Table 12.2 Critical Values for the Dickey–Fuller Test

Model	1%	5%	10%
$\Delta y_t = \gamma y_{t-1} + v_t$	−2.56	−1.94	−1.62
$\Delta y_t = \alpha + \gamma y_{t-1} + v_t$	−3.43	−2.86	−2.57
$\Delta y_t = \alpha + \lambda t + \gamma y_{t-1} + v_t$	−3.96	−3.41	−3.13
Standard critical values	−2.33	−1.65	−1.28

Note: These critical values are taken from R. Davidson and J. G. MacKinnon (1993), *Estimation and Inference in Econometrics*, New York: Oxford University Press, p. 708.

and nonstationarity are expressed in terms of γ in the same way and the test critical values are the same as those for the Dickey–Fuller test shown in Table 12.2. When $\gamma = 0$, in addition to saying that the series is nonstationary, we also say the series has a **unit root**. In practice, we always use the augmented Dickey–Fuller test (rather than the nonaugmented version) to ensure the errors are uncorrelated.

12.3.5 The Dickey Fuller Testing Procedures

Up to now, we have discussed a number of stationary and nonstationary processes as well as three Dickey-Fuller tests. How do we go about deciding which test to use? To understand the rationale for what we suggest, it is useful to first take a look at the design of the unit root tests.

The critical values for the three tests shown in Table 12.2 were derived from the following simulations:

- true process; $y_t = y_{t-1} + v_t, v_t \sim N(0, \sigma^2)$, test equation: $y_t = \rho y_{t-1} + v_t$
- true process; $y_t = y_{t-1} + v_t, v_t \sim N(0, \sigma^2)$, test equation: $y_t = \alpha + \rho y_{t-1} + v_t$
- true process; $y_t = \delta + y_{t-1} + v_t, v_t \sim N(0, \sigma^2)$, test equation: $y_t = \alpha + \rho y_{t-1} + \lambda t + v_t$

Now take a look at Table 12.3. Column one shows the stationary autoregressive models covered in Section 12.1.1, and column two shows the corresponding nonstationary processes when $\rho = 1$. As we can see the processes in column two correspond to the true processes underlying the Dickey-Fuller tests described in column three while the processes in column one are the test equations.

Table 12.3 AR processes and the Dickey-Fuller Tests

AR processes: $\lvert\rho\rvert < 1$	Setting $\rho = 1$	Dickey Fuller Tests
$y_t = \rho y_{t-1} + u_t$	$y_t = y_{t-1} + u_t$	Test with no constant and no trend
$y_t = \alpha + \rho y_{t-1} + v_t$ $\alpha = \mu(1-\rho)$	$y_t = y_{t-1} + v_t$ $\alpha = 0$	Test with constant and no trend
$y_t = \alpha + \rho y_{t-1} + \lambda t + v_t$ $\alpha = (\mu(1-\rho) + \rho\delta)$ $\lambda = \delta(1-\rho)$	$y_t = \delta + y_{t-1} + v_t$ $\alpha = \delta$ $\lambda = 0$	Test with constant and trend

This then suggests the following Dickey-Fuller testing procedure. First plot the time series of the variable and select a suitable Dickey-Fuller test based on a visual inspection of the plot.

- If the series appears to be wandering or fluctuating around a sample average of zero (see for example Figure 12.2(a) or Figure 12.2(d) with mean around zero), use test equation (12.5a).
- If the series appears to be wandering or fluctuating around a sample average which is nonzero (see for example Figure 12.2(b) or Figure 12.2(d) with a non-zero mean), use test equation (12.5b).
- If the series appears to be wandering or fluctuating around a linear trend (see, for example, Figure 12.5(c) or Figure 12.2(e)), use test equation (12.5c).

Second, proceed with one of the unit root tests described in Sections 12.3.1 to 12.3.3, bearing in mind that it is important to choose the correct critical values as they depend upon the test equation estimated, which, in turn, depends on the absence or presence of the constant and trend terms.

12.3.6 The Dickey–Fuller Tests: An Example

As an example, consider the two interest rate series—the federal funds rate (F_t) and the three-year bond rate (B_t)—plotted in Figure 12.1(e) and 12.1(g), respectively. Both series exhibit wandering behavior, so we suspect that they may be nonstationary variables. When performing Dickey–Fuller tests, we need to decide whether to use (12.5a) with no constant, or (12.5b) that includes a constant term, or (12.5c) that includes a constant and a deterministic time trend t. As suggested earlier, (12.5b) is the appropriate test equation because the series fluctuate around a nonzero mean. We also have to decide on how many lagged difference terms to include on the right-hand side of the equation. Following procedures described in Sections 9.3 and 9.4, we find that the inclusion of one lagged difference term is sufficient to eliminate autocorrelation in the residuals in both cases. The results from estimating the resulting equations are

$$\begin{aligned}\widehat{\Delta F_t} &= 0.173 - 0.045F_{t-1} + 0.561\Delta F_{t-1}\\ (tau) &\qquad\quad (-2.505)\end{aligned}$$

$$\begin{aligned}\widehat{\Delta B_t} &= 0.237 - 0.056B_{t-1} + 0.237\Delta B_{t-1}\\ (tau) &\qquad\quad (-2.703)\end{aligned}$$

The *tau* value (τ) for the federal funds rate is -2.505, and the 5% critical value for *tau* (τ_c) is -2.86. Again, recall that to carry out this one-tail test of significance, we reject the null hypothesis of nonstationarity if $\tau \leq \tau_c$. If $\tau > \tau_c$ then we do not reject the null hypothesis that the series is nonstationary. In this case, since $-2.505 > -2.86$, we do not reject the null hypothesis that the series is nonstationary. Similarly, the *tau* value for the bond rate is greater than the 5% critical value of -2.86 and again we do not reject the null hypothesis that the series is nonstationary. Expressed another way, there is insufficient evidence to suggest F_t and B_t are stationary.

12.3.7 Order of Integration

Up to this stage, we have discussed only whether a series is stationary or nonstationary. We can take the analysis another step forward and consider a concept called the ‘‘order of integration.'' Recall that if y_t follows a random walk, then $\gamma = 0$ and the first difference of y_t becomes

$$\Delta y_t = y_t - y_{t-1} = v_t$$

An interesting feature of the series $\Delta y_t = y_t - y_{t-1}$ is that it is stationary since v_t, being an independent $(0, \sigma_v^2)$ random variable, is stationary. Series like y_t, which can be made stationary by taking the first difference, are said to be **integrated of order one**, and denoted as **I(1)**. Stationary series are said to be integrated of order zero, **I(0)**. In general, the order of integration of a series is the minimum number of times it must be differenced to make it stationary.

For example, to determine the order of integration of F and B, we then ask the next question: is the first difference of the federal funds rate $(\Delta F_t = F_t - F_{t-1})$ stationary? Is the first difference of the bond rate $(\Delta B_t = B_t - B_{t-1})$ stationary? Their plots, in Figure 12.1(f) and 12.1(h), seem to suggest that they are stationary.

The results of the Dickey–Fuller test for a random walk applied to the first differences are given below:

$$\begin{array}{rl}\widehat{\Delta(\Delta F)_t} = & -0.447(\Delta F)_{t-1} \\ (tau) & (-5.487)\end{array}$$

$$\begin{array}{rl}\widehat{\Delta(\Delta B)_t} = & -0.701(\Delta B)_{t-1} \\ (tau) & (-7.662)\end{array}$$

where $\Delta(\Delta F)_t = \Delta F_t - \Delta F_{t-1}$ and $\Delta(\Delta B)_t = \Delta B_t - \Delta B_{t-1}$. Note that the null hypotheses are that the variables ΔF and ΔB are not stationary. Also, because the series ΔF and ΔB appear to fluctuate around zero, we use the test equation without the intercept term. Based on the large negative value of the *tau* statistic $(-5.487 < -1.94)$, we reject the null hypothesis that ΔF_t is nonstationary and accept the alternative that it is stationary. We similarly conclude that ΔB_t is stationary $(-7.662 < -1.94)$.

This result implies that while the level of the federal funds rate (F_t) is nonstationary, its first difference (ΔF_t) is stationary. We say that the series F_t is I(1) because it had to be differenced once to make it stationary [ΔF_t is I(0)]. Similarly we have also shown that the bond rate (B_t) is integrated of order one. In the next section we investigate the implications of these results for regression modeling.

12.4 Cointegration

As a general rule, nonstationary time-series variables should not be used in regression models, to avoid the problem of spurious regression. However, there is an exception to this rule. If y_t and x_t are nonstationary I(1) variables, then we expect their difference, or any linear combination of them, such as $e_t = y_t - \beta_1 - \beta_2 x_t$,[3] to be I(1) as well. However, there is an important case when $e_t = y_t - \beta_1 - \beta_2 x_t$ is a stationary I(0) process. In this case y_t and x_t are said to be **cointegrated**. Cointegration implies that y_t and x_t share similar stochastic trends, and, since the difference e_t is stationary, they never diverge too far from each other.

A natural way to test whether y_t and x_t are cointegrated is to test whether the errors $e_t = y_t - \beta_1 - \beta_2 x_t$ are stationary. Since we cannot observe e_t, we test the stationarity of the

[3] A linear combination of x and y is a new variable $z = a_0 + a_1 x + a_2 y$. Here we set the constants $a_0 = -\beta_1$, $a_1 = -\beta_2$, and $a_2 = 1$, and call z the series e.

Table 12.4 Critical Values for the Cointegration Test

Regression model	1%	5%	10%
(1) $y_t = \beta x_t + e_t$	−3.39	−2.76	−2.45
(2) $y_t = \beta_1 + \beta_2 x_t + e_t$	−3.96	−3.37	−3.07
(3) $y_t = \beta_1 + \delta t + \beta_2 x_t + e_t$	−3.98	−3.42	−3.13

Note: These critical values are taken from J. Hamilton (1994), *Time Series Analysis*, Princeton University Press, p. 766.

least squares residuals, $\hat{e}_t = y_t - b_1 - b_2 x_t$ using a Dickey–Fuller test. The test for cointegration is effectively a test of the stationarity of the residuals. If the residuals are stationary, then y_t and x_t are said to be cointegrated; if the residuals are nonstationary, then y_t and x_t are not cointegrated, and any apparent regression relationship between them is said to be spurious.

The test for stationarity of the residuals is based on the test equation

$$\Delta\hat{e}_t = \gamma\hat{e}_{t-1} + v_t \tag{12.7}$$

where $\Delta\hat{e}_t = \hat{e}_t - \hat{e}_{t-1}$. As before, we examine the t (or *tau*) statistic for the estimated slope coefficient. Note that the regression has no constant term because the mean of the regression residuals is zero. Also, since we are basing this test upon **estimated** values of the residuals, the critical values will be different from those in Table 12.2. The proper critical values for a test of cointegration are given in Table 12.4. The test equation can also include extra terms like $\Delta\hat{e}_{t-1}, \Delta\hat{e}_{t-2}, \ldots$ on the right-hand side if they are needed to eliminate autocorrelation in v_t.

There are three sets of critical values. Which set we use depends on whether the residuals $\hat{e}_t$ are derived from a regression equation without a constant term [like (12.8a)] or a regression equation with a constant term [like (12.8b)], or a regression equation with a constant and a time trend [like (12.8c)].

$$\text{Equation 1:}\quad \hat{e}_t = y_t - b x_t \tag{12.8a}$$

$$\text{Equation 2:}\quad \hat{e}_t = y_t - b_2 x_t - b_1 \tag{12.8b}$$

$$\text{Equation 3:}\quad \hat{e}_t = y_t - b_2 x_t - b_1 - \hat{\delta} t \tag{12.8c}$$

12.4.1 An Example of a Cointegration Test

To illustrate, let us test whether $y_t = B_t$ and $x_t = F_t$, as plotted in Figure 12.1(e) and 12.1(g), are cointegrated. We have already shown that both series are nonstationary. The estimated least squares regression between these variables is

$$\begin{aligned} &\hat{B}_t = 1.140 + 0.914 F_t, \quad R^2 = 0.881 \\ &(t) \quad (6.548)(29.421) \end{aligned} \tag{12.9}$$

and the unit root test for stationarity in the estimated residuals ($\hat{e}_t = B_t - 1.140 - 0.914 F_t$) is

$$\begin{aligned} &\Delta\hat{e}_t = -0.225\hat{e}_{t-1} + 0.254\Delta\hat{e}_{t-1} \\ &(tau) \quad (-4.196) \end{aligned}$$

Note that this is the augmented Dickey–Fuller version of the test with one lagged term Δe_{t-1} to correct for autocorrelation. Since there is a constant term in (12.9), we use the equation (2) critical values in Table 12.4.

The null and alternative hypotheses in the test for cointegration are

$$H_0: \text{the series are not cointegrated} \Leftrightarrow \text{residuals are nonstationary}$$

$$H_1: \text{the series are cointegrated} \Leftrightarrow \text{residuals are stationary}$$

Similar to the one-tail unit root tests, we reject the null hypothesis of no cointegration if $\tau \leq \tau_c$, and we do not reject the null hypothesis that the series are not cointegrated if $\tau > \tau_c$. The *tau* statistic in this case is -4.196 which is less than the critical value -3.37 at the 5% level of significance. Thus, we reject the null hypothesis that the least squares residuals are nonstationary and conclude that they are stationary. This implies that the bond rate and the federal funds rate are cointegrated. In other words, there is a fundamental relationship between these two variables (the estimated regression relationship between them is valid and not spurious) and the estimated values of the intercept and slope are 1.140 and 0.914, respectively.

The result—that the federal funds and bond rates are cointegrated—has major economic implications! It means that when the Federal Reserve implements monetary policy by changing the federal funds rate, the bond rate will also change thereby ensuring that the effects of monetary policy are transmitted to the rest of the economy. In contrast, the effectiveness of monetary policy would be severely hampered if the bond and federal funds rates were spuriously related as this implies that their movements, fundamentally, have little to do with each other.

12.4.2 The Error Correction Model

In the previous section, we discussed the concept of cointegration as the relationship between I(1) variables such that the residuals are I(0). A relationship between I(1) variables is also often referred to as a long run relationship while a relationship between I(0) variables is often referred to as a short run relationship. In this section, we describe a dynamic relationship between I(0) variables, which embeds a cointegrating relationship, known as the short-run error correction model.

As discussed in Chapter 9, when one is working with time-series data, it is quite common, and in fact, is quite important to allow for dynamic effects. To derive the error correction model requires a bit of algebra, but we shall persevere as this model offers a coherent way to combine the long- and short-run effects.

Let us start with a general model that contains lags of y and x, namely the autoregressive distributed lag (ARDL) model introduced in Chapter 9, except that now the variables are nonstationary:

$$y_t = \delta + \theta_1 y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + v_t$$

For simplicity, we shall consider lags up to order one, but the following analysis holds for any order of lags. Now recognize that if y and x are cointegrated, it means that there is a long-run relationship between them. To derive this exact relationship, we set $y_t = y_{t-1} = y$, $x_t = x_{t-1} = x$ and $v_t = 0$ and then, imposing this concept in the ARDL, we obtain

$$y(1 - \theta_1) = \delta + (\delta_0 + \delta_1)x$$

This equation can be rewritten as $y = \beta_1 + \beta_2 x$ where $\beta_1 = \delta/(1 - \theta_1)$ and $\beta_2 = (\delta_0 + \delta_1)/(1 - \theta_1)$. To repeat, we have now derived the implied cointegrating relationship between y and x; alternatively, we have derived the long-run relationship that holds between the two I(1) variables.

We will now manipulate the ARDL to see how it embeds the cointegrating relation. First, add the term, $-y_{t-1}$, to both sides of the equation:

$$y_t - y_{t-1} = \delta + (\theta_1 - 1)y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + v_t.$$

Second, add the term $-\delta_0 x_{t-1} + \delta_0 x_{t-1}$ to the right-hand side to obtain

$$\Delta y_t = \delta + (\theta_1 - 1)y_{t-1} + \delta_0(x_t - x_{t-1}) + (\delta_0 + \delta_1)x_{t-1} + v_t$$

where $\Delta y_t = y_t - y_{t-1}$. If we then manipulate the equation to look like

$$\Delta y_t = (\theta_1 - 1)\left(\frac{\delta}{(\theta_1 - 1)} + y_{t-1} + \frac{(\delta_0 + \delta_1)}{(\theta_1 - 1)}x_{t-1}\right) + \delta_0 \Delta x_t + v_t$$

where $\Delta x_t = x_t - x_{t-1}$, and do a bit more tidying, using the definitions β_1 and β_2, we get

$$\Delta y_t = -\alpha(y_{t-1} - \beta_1 - \beta_2 x_{t-1}) + \delta_0 \Delta x_t + v_t \tag{12.10}$$

where $\alpha = (1 - \theta_1)$. As you can see, the expression in parenthesis is the cointegrating relationship. In other words, we have embedded the cointegrating relationship between y and x in a general ARDL framework.

Equation (12.10) is called an error correction equation because (a) the expression $(y_{t-1} - \beta_1 - \beta_1 x_{t-1})$ shows the deviation of y_{t-1} from its long run value, $\beta_1 + \beta_2 x_{t-1}$—in other words, the "error" in the previous period—and (b) the term $(\theta_1 - 1)$ shows the "correction" of Δy_t to the "error." More specifically, if the error in the previous period is positive so that $y_{t-1} > (\beta_0 + \beta_1 x_{t-1})$, then y_t should fall and Δy_t should be negative; conversely, if the error in the previous period is negative so that $y_{t-1} < (\beta_0 + \beta_1 x_{t-1})$, then y_t should rise and Δy_t should be positive. This means that if a cointegrating relationship between y and x exists, so that adjustments always work to "error-correct," then empirically we should also find that $(1 - \theta_1) > 0$, which implies that $\theta_1 < 1$. If there is no evidence of cointegration between the variables, then the term θ_1 would be insignificant.

The error correction model is a very popular model because it allows for the existence of an underlying or fundamental link between variables (the long-run relationship) as well as for short-run adjustments (i.e. changes) between variables, including adjustments to achieve the cointegrating relationship. It also shows that we can work with I(1) variables (y_{t-1}, x_{t-1}) and I(0) variables (Δy_t, Δx_t) in the same equation provided that (y, x) are cointegrated, meaning that the term $(y_{t-1} - \beta_0 - \beta_1 x_{t-1})$ contains stationary residuals. In fact, this formulation can also be used to test for cointegration between y and x.

To illustrate, consider our earlier example of the bond and federal funds rates. The result from estimating (12.10) using nonlinear least squares is

$$\begin{array}{llll} \Delta\hat{B}_t = -0.142(B_{t-1} - 1.429 - 0.777F_{t-1}) & + 0.842\Delta F_t & - 0.327\Delta F_{t-1} \\ (t) \qquad (2.857) & \qquad (9.387) & \qquad (3.855) \end{array}$$

Note first that we need two lags $(\Delta F_t, \Delta F_{t-1})$ to ensure that the residuals are purged of all serial correlation effects. Second, note that the estimate $\hat{\theta}_1 = -0.142 + 1 = 0.858$ is less than one, as expected.

We can now generate the estimated residuals:

$$\hat{e}_{t-1} = (B_{t-1} - 1.429 - 0.777F_{t-1})$$

The result from applying the ADF test for stationarity is

$$\begin{aligned}&\Delta\hat{e}_t = -0.169\hat{e}_{t-1} + 0.180\Delta\hat{e}_{t-1}\\ &(t) \quad (-3.929)\end{aligned}$$

As before, the null is that (B, F) are not cointegrated. Since the cointegrating relationship includes a constant term, the critical value is -3.37. Comparing the calculated value (-3.929) with the critical value, we reject the null hypothesis and conclude that (B, F) are cointegrated.

12.5 Regression When There Is No Cointegration

Thus far, we have shown that regression with I(1) variables is acceptable providing those variables are cointegrated, allowing us to avoid the problem of spurious results. We also know that regression with stationary I(0) variables, that we studied in Chapter 9, is acceptable. What happens when there is no cointegration between I(1) variables? In this case, the sensible thing to do is to convert the nonstationary series to stationary series and to use the techniques discussed in Chapter 9 to estimate dynamic relationships between the stationary variables. However, we stress that this step should be taken only when we fail to find cointegration between the I(1) variables. Regression with cointegrated I(1) variables makes the least squares estimator "super-consistent"[4] and, moreover, is economically useful to establish relationships between the levels of economic variables.

How we convert nonstationary series to stationary series, and the kind of model we estimate, depend on whether the variables are **difference stationary** or **trend stationary**. In the former case, we convert the nonstationary series to its stationary counterpart by taking first differences. In the latter case, we convert the nonstationary series to its stationary counterpart by de-trending. We now explore these issues.

12.5.1 First Difference Stationary

Consider a variable y_t that behaves like the random walk model:

$$y_t = y_{t-1} + v_t$$

This is a nonstationary series with a "stochastic" trend, but it can be rendered stationary by taking the first difference:

$$\Delta y_t = y_t - y_{t-1} = v_t$$

The variable y_t is said to be a **first difference stationary** series. Recall that this means that y is said to be integrated of order 1. Now suppose that Dickey–Fuller tests reveal that two variables, y and x, that you would like to relate in a regression, are first difference stationary,

[4]Consistency means that as $T \rightarrow \infty$ the least squares estimator converges to the true parameter value. See Appendix 5B. Super-consistency means that it converges to the true value at a faster rate.

I(1), and not cointegrated. Then, a suitable regression involving only stationary variables is one that relates changes in y to changes in x, with relevant lags included, and no intercept. For example, using one lagged Δy_t and a current and lagged Δx_t, we have

$$\Delta y_t = \theta \Delta y_{t-1} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + e_t \tag{12.11a}$$

Now consider a series y_t that behaves like a random walk with drift,

$$y_t = \alpha + y_{t-1} + v_t$$

and note that y can be rendered stationary by taking the first difference:

$$\Delta y_t = \alpha + v_t$$

The variable y_t is also said to be a **first difference stationary** series, even though it is stationary around a constant term. Now suppose again that y and x are I(1) and not cointegrated. Then an example of a suitable regression equation, again involving stationary variables, is obtained by adding a constant to (12.11a). That is,

$$\Delta y_t = \alpha + \theta \Delta y_{t-1} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + e_t \tag{12.11b}$$

In line with Section 9.7, the models in (12.11a) and (12.11b) are autoregressive distributed lag models with first-differenced variables. In general, since there is often doubt about the role of the constant term, the usual practice is to include an intercept term in the regression.

12.5.2 Trend Stationary

Consider a model with a constant term, a trend term, and a stationary error term:

$$y_t = \alpha + \delta t + v_t$$

The variable y_t is said to be **trend stationary** because it can be made stationary by removing the effect of the deterministic (constant and trend) components

$$y_t - \alpha - \delta t = v_t$$

A series like this is, strictly speaking, not an I(1) variable, but is described as stationary around a deterministic trend. Thus, if y and x are two trend-stationary variables, a possible autoregressive distributed lag model is

$$y_t^* = \theta y_{t-1}^* + \beta_0 x_t^* + \beta_1 x_{t-1}^* + e_t \tag{12.12}$$

where $y_t^* = y_t - \alpha_1 - \delta_1 t$ and $x_t^* = x_t - \alpha_2 - \delta_2 t$ are the de-trended data (the coefficients (α_1, δ_1) and (α_2, δ_2) can be estimated by least squares).

As an alternative to using the de-trended data for estimation, a constant term and a trend term can be included directly in the equation. For example, by substituting y_t^* and x_t^* into (12.12), it can be shown that estimating (12.12) is equivalent to estimating

$$y_t = \alpha + \delta t + \theta y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + e_t$$

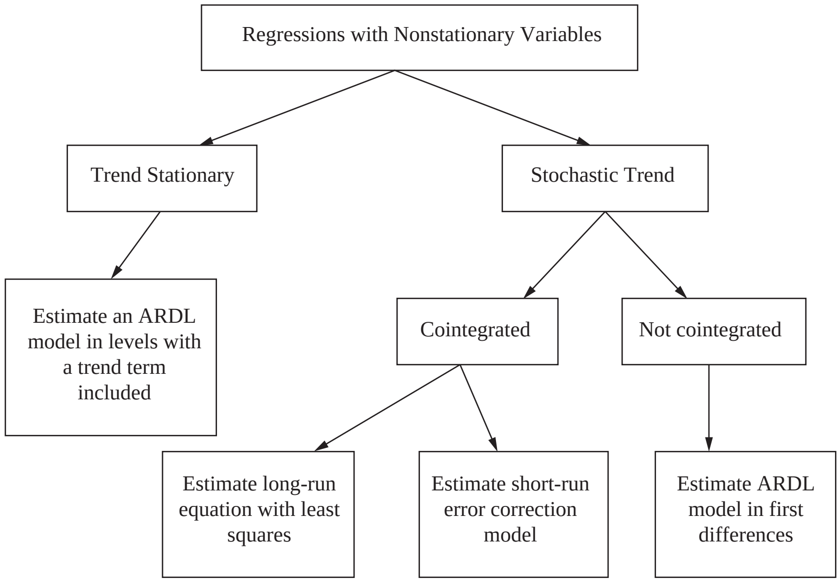

FIGURE 12.4 Regression with time-series data: nonstationary variables.

where $\alpha = \alpha_1(1 - \theta_1) - \alpha_2(\beta_0 + \beta_1) + \theta_1\delta_1 + \beta_1\delta_2$ and $\delta = \delta_1(1 - \theta_1) - \delta_2(\beta_0 + \beta_1)$. In practice, this is usually the preferred option as it is relatively more straightforward.

12.5.3 Summary

- If variables are stationary, or I(1) and cointegrated, we can estimate a regression relationship between the levels of those variables without fear of encountering a spurious regression. In the later case, we can do this by estimating a least squares equation between the I(1) variables or by estimating a nonlinear least squares error correction model which embeds the I(1) variables.
- If the variables are I(1) and not cointegrated, we need to estimate a relationship in first differences, with or without the constant term.
- If they are trend stationary, we can either de-trend the series first and then perform regression analysis with the stationary (de-trended) variables or, alternatively, estimate a regression relationship that includes a trend variable. The latter alternative is typically applied.

These options are shown in Figure 12.4.

12.6 Exercises

12.6.1 Problems

12.1 (a) Consider an AR(1) model

$$y_t = \rho y_{t-1} + v_t, \quad |\rho| < 1$$

Rewrite y as a function of lagged errors. (*Hint*: perform recursive substitution.) What is the mean and variance of y? What is the covariance between y_t and y_{t-2}?

(b) Consider a random walk model

$$y_t = y_{t-1} + v_t$$

Rewrite y as a function of lagged errors. What is the mean and variance of y? What is the covariance between y_t and y_{t-2}?

12.2 Figure 12.5 (data file *unit.dat*) shows plots of four time series. Since W and Y appear to be fluctuating around a nonzero mean, a Dickey–Fuller test 2 (with constant but no trend) was performed on these variables. Since X and Z appear to be fluctuating around a trend, a Dickey–Fuller test 3 (with constant and trend) was performed for these two variables. The results are shown below.

$$\widehat{\Delta W_t} = 0.757 - 0.091 W_{t-1}$$
$$(tau) \qquad (-3.178)$$

$$\widehat{\Delta Y_t} = 0.031 - 0.039 Y_{t-1}$$
$$(tau) \qquad (-1.975)$$

$$\widehat{\Delta X_t} = 0.782 - 0.092 X_{t-1} + 0.009t$$
$$(tau) \qquad (-3.099)$$

$$\widehat{\Delta Z_t} = 0.332 - 0.036 Z_{t-1} + 0.005t$$
$$(tau) \qquad (-1.913)$$

Which series are stationary, and which are nonstationary?

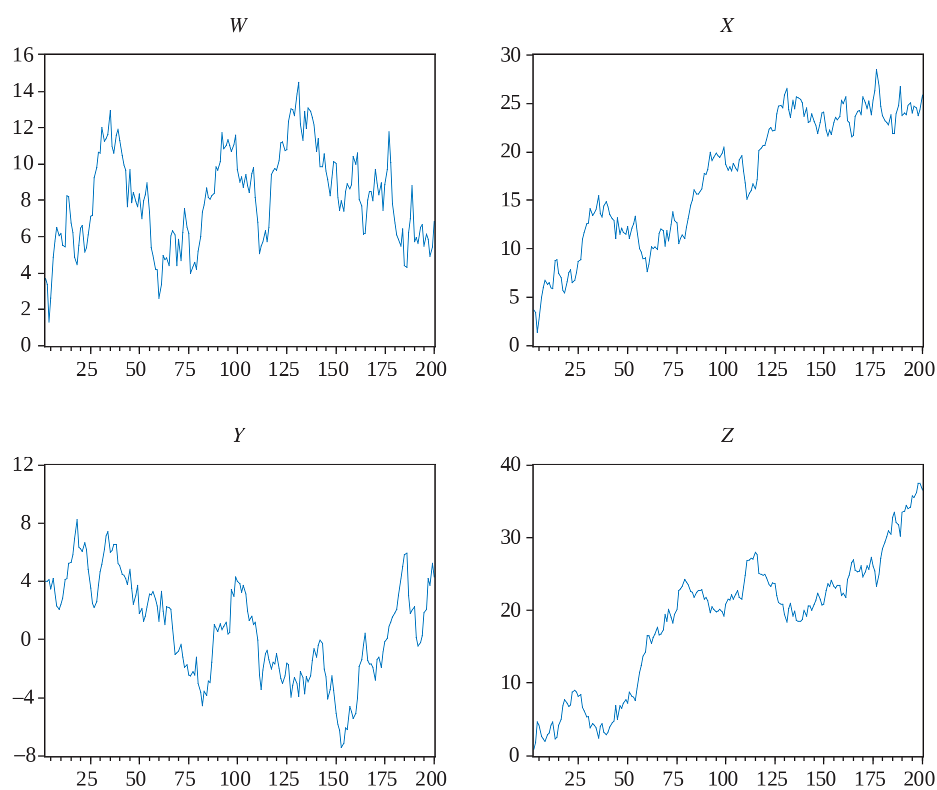

FIGURE 12.5 Time series for Exercise 12.2.

12.3 A time series process of the form $y_t = \alpha + y_{t-1} + v_t, v_t \sim N(0, \sigma^2)$ can be rearranged as $y_t - y_{t-1} = \Delta y_t = \alpha + v_t$. This shows that y_t is integrated of order one, since its first difference is stationary. Show that a time series of the form $y_t = 2y_{t-1} - y_{t-2} + \alpha + v_t$ is integrated of order two.

12.6.2 Computer Exercises

12.4 The data file *oil.dat* contains 88 annual observations on the price of oil (in 1967 constant dollars) for the period 1883–1970.
(a) Plot the data. Do the data look stationary, or nonstationary?
(b) Use a unit root test to demonstrate that the series is stationary.
(c) What do you conclude about the order of integration of this series?

12.5 The data file *bond.dat* contains 102 monthly observations on AA railroad bond yields for the period January 1968 to June 1976.
(a) Plot the data. Do railroad bond yields appear stationary, or nonstationary?
(b) Use a unit root test to demonstrate that the series is nonstationary.
(c) Find the first difference of the bond yield series and test for stationarity.
(d) What do you conclude about the order of integration of this series?

12.6 The data file *oz.dat* contains quarterly data on disposable income and consumption in Australia from 1985:1 to 2005:2.
(a) Test each of these series for stationarity.
(b) What do you conclude about the "order of integration" of each of these series?
(c) Is consumption cointegrated with, or spuriously related to, disposable income?

12.7 The data file *texas.dat* contains 57 quarterly observations on the real price of oil (*RPO*), Texas nonagricultural employment (*TXNAG*), and nonagricultural employment in the rest of the United States (*USNAG*). The data cover the period 1974Q1 through 1988Q1 and were used in a study by Fomby and Hirschberg [T. B. Fomby and J. G. Hirschberg, "Texas in Transition: Dependence on Oil and the National Economy," *Federal Reserve Bank of Dallas Economic Review*, January 1989, 11–28].
(a) Show that the **levels** of the variables *TXNAG* and *USNAG* are nonstationary variables.
(b) At what significance level do you conclude that the **changes** $DTX = TXNAG - TXNAG(-1)$ and $DUS = USNAG - USNAG(-1)$ are stationary variables?
(c) Are the nonstationary variables *TXNAG* and *USNAG* cointegrated, or spuriously related?
(d) Are the stationary variables *DTX* and *DUS* related?
(e) What is the difference between (d) and (c)?

12.8 The data file *usa.dat* contains the data shown in Figure 12.1. Consider the two time series, real GDP and the inflation rate.
(a) Are the series stationary, or nonstationary? Which Dickey–Fuller test (no constant, no trend; with constant, no trend; or with constant and with trend) did you use?
(b) What do you conclude about the order of integration of these series?
(c) Forecast GDP and inflation for 2010:1.

12.9 The data file *canada.dat* contains monthly Canadian/U.S. exchange rates for the period 1971:01 to 2006:12. Split the observations into two sample periods—a 1971:01–1987:12 sample period and a 1988:01–2006:12 sample period.
(a) Perform a unit root test on the data for each sample period. Which Dickey–Fuller test did you use?
(b) Are the results for the two sample periods consistent?

(c) Perform a unit root test for the full sample 1971:01–2006:12. What is the order of integration of the data?

12.10 The data file *csi.dat* contains the Consumer Sentiment Index (CSI), produced by the University of Michigan for the sample period 1978:01–2006:12.
(a) Perform all three Dickey–Fuller tests. Are the results consistent? If not, why not?
(b) Based on a graphical inspection of the data, which test should you have used?
(c) Does the CSI suggest that consumers ''remember'' and ''retain'' news information for a short time, or for a long time?

12.11 The data file *mexico.dat* contains real GDP for Mexico and the Unites States from the first quarter of 1980 to the third quarter of 2006. Both series have been standardized so that the average value in 2,000 is 100.
(a) Perform the test for cointegration between Mexico and the Unites States for all three test equations in (12.8). Are the results consistent?
(b) The theory of convergence in economic growth suggests the two GDPs should be proportional and cointegrated. That is, there should be a cointegrating relationship that does not contain an intercept or a trend. Do your results support this theory?
(c) If the variables are not cointegrated, what should you do if you are interested in testing the relationship between Mexico and the United States?

12.12 The file *inter2.dat* contains 300 observations of a generated I(2) process shown in Figure 12.6 below. Show that the variable called *inter2* is indeed an I(2) variable by conducting a number of unit root tests—first on the level of the data, then on the first difference and finally on the second difference.

12.13 Prices around the world tend to move together. The data file *ukpi.dat* contains information about the price indices in the United Kingdom and in the Euro Area (the United Kingdom is a member of the European Union, but not a member of the single European currency zone) for the period 1996:1–2009:12.
(a) Plot the data. Are the series I(1) or I(0)?
(b) Are prices in the UK and in the Euro Area cointegrated, or spuriously related? Use both the least squares and the error correction method to test this proposition.

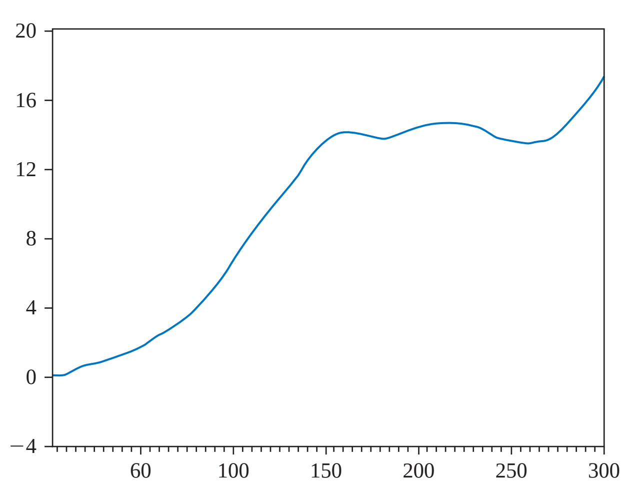

FIGURE 12.6 A generated I(2) process.