21

TIME SERIES ECONOMETRICS: SOME BASIC CONCEPTS

We noted in Chapter 1 that one of the important types of data used in empirical analysis is **time series** data. In this and the following chapter we take a closer look at such data not only because of the frequency with which they are used in practice but also because they pose several challenges to econometricians and practitioners.

First, empirical work based on time series data assumes that the underlying time series is **stationary.** Although we have discussed the concept of stationarity intuitively in Chapter 1, we discuss it more fully in this chapter. More specifically, we will try to find out what stationarity means and why one should worry about it.

Second, in Chapter 12, on autocorrelation, we discussed several causes of autocorrelation. Sometimes autocorrelation results because the underlying time series is nonstationary.

Third, in regressing a time series variable on another time series variable(s), one often obtains a very high R^2 (in excess of 0.9) even though there is no meaningful relationship between the two variables. Sometimes we expect no relationship between two variables, yet a regression of one on the other variable often shows a significant relationship. This situation exemplifies the problem of **spurious,** or **nonsense, regression,** whose nature will be explored shortly. It is therefore very important to find out if the relationship between economic variables is spurious or nonsensical. We will see in this chapter how spurious regressions can arise if time series are not stationary.

Fourth, some financial time series, such as stock prices, exhibit what is known as the **random walk phenomenon.** This means the best prediction of the price of a stock, say IBM, tomorrow is equal to its price today plus a purely random shock (or error term). If this were in fact the case, forecasting asset prices would be a futile exercise.

Fifth, regression models involving time series data are often used for forecasting. In view of the preceding discussion, we would like to know if such forecasting is valid if the underlying time series are not stationary.

Finally, causality tests of Granger and Sims that we discussed in Chapter 17 assume that the time series involved in analysis are stationary. Therefore, tests of stationarity should precede tests of causality.

At the outset a disclaimer is in order. The topic of time series analysis is so vast and evolving and some of the mathematics underlying the various techniques of time series analysis is so involved that the best we hope to achieve in an introductory text like this is to give the reader a glimpse of some of the fundamental concepts of time series analysis. For those who want to pursue this topic further, we provide references.[1]

21.1 A LOOK AT SELECTED U.S. ECONOMIC TIME SERIES

To set the stage, and to give the reader a feel for the somewhat esoteric concepts of time series analysis to be developed in this chapter, it might be useful to consider several U.S. economic time series of general interest. The time series we consider are: (1) GDP (gross domestic product), (2) PDI (personal disposable income), (3) PCE (personal consumption expenditure), (4) profits (corporate profits after tax), and (5) dividends (net corporate dividend); all data are in billions of 1987 dollars and are for the quarterly periods of 1970–1991, for a total of 88 quarterly observations. The raw data are given in Table 21.1.

[1]At the introductory level, these references may be helpful: Gary Koop, *Analysis of Economic Data,* John Wiley & Sons, New York, 2000; Jeff B. Cromwell, Walter C. Labys, and Michel Terraza, *Univariate Tests for Time Series Models,* Sage Publications, California, Ansbury Park, 1994; Jeff B. Cromwell, Michael H. Hannan, Walter C. Labys, and Michel Terraza, *Multivariate Tests for Time Series Models,* Sage Publications, California, Ansbury Park, 1994; H. R. Seddighi, K. A. Lawler, and A. V. Katos, *Econometrics: A Practical Approach,* Routledge, New York, 2000. At the intermediate level, see Walter Enders, *Applied Econometric Time Series,* John Wiley & Sons, New York, 1995; Kerry Patterson, *An Introduction to Applied Econometrics: A Time Series Approach,* St. Martin's Press, New York, 2000; T. C. Mills, *The Econometric Modelling of Financial Time Series,* 2d ed., Cambridge University Press, New York, 1999; Marno Verbeek, *A Guide to Modern Econometrics,* John Wiley & Sons, New York, 2000; Wojciech W. Charemza and Derek F. Deadman, *New Directions in Econometric Practice: General to Specific Modelling and Vector Autoregression,* 2d ed., Edward Elgar Publisher, New York, 1997. At the advanced level, see Hamilton, J. D., *Time Series Analysis,* Princeton University Press, Princeton, N.J., 1994, and G. S. Maddala and In-Moo Kim, *Unit Roots, Cointegration, and Structural Change,* Cambridge University Press, 1998. At the applied level, see B. Bhaskara Rao, ed., *Cointegration for the Applied Economist,* St. Martin's Press, New York, 1994, and Chandan Mukherjee, Howard White, and Marc Wuyts, *Econometrics and Data Analysis for Developing Countries,* Routledge, New York, 1998.

TABLE 21.1 MACROECONOMICS DATA, UNITED STATES, 1970–I TO 1991–IV

Quarter	GDP	PDI	PCE	Profits	Dividend	Quarter	GDP	PDI	PCE	Profits	Dividend
1970–I	2872.8	1990.6	1800.5	44.7	24.5	1981–I	3860.5	2783.7	2475.5	159.5	64.0
1970–II	2860.3	2020.1	1807.5	44.4	23.9	1981–II	3844.4	2776.7	2476.1	143.7	68.4
1970–III	2896.6	2045.3	1824.7	44.9	23.3	1981–III	3864.5	2814.1	2487.4	147.6	71.9
1970–IV	2873.7	2045.2	1821.2	42.1	23.1	1981–IV	3803.1	2808.8	2468.6	140.3	72.4
1971–I	2942.9	2073.9	1849.9	48.8	23.8	1982–I	3756.1	2795.0	2484.0	114.4	70.0
1971–II	2947.4	2098.0	1863.5	50.7	23.7	1982–II	3771.1	2824.8	2488.9	114.0	68.4
1971–III	2966.0	2106.6	1876.9	54.2	23.8	1982–III	3754.4	2829.0	2502.5	114.6	69.2
1971–IV	2980.8	2121.1	1904.6	55.7	23.7	1982–IV	3759.6	2832.6	2539.3	109.9	72.5
1972–I	3037.3	2129.7	1929.3	59.4	25.0	1983–I	3783.5	2843.6	2556.5	113.6	77.0
1972–II	3089.7	2149.1	1963.3	60.1	25.5	1983–II	3886.5	2867.0	2604.0	133.0	80.5
1972–III	3125.8	2193.9	1989.1	62.8	26.1	1983–III	3944.4	2903.0	2639.0	145.7	83.1
1972–IV	3175.5	2272.0	2032.1	68.3	26.5	1983–IV	4012.1	2960.6	2678.2	141.6	84.2
1973–I	3253.3	2300.7	2063.9	79.1	27.0	1984–I	4089.5	3033.2	2703.8	155.1	83.3
1973–II	3267.6	2315.2	2062.0	81.2	27.8	1984–II	4144.0	3065.9	2741.1	152.6	82.2
1973–III	3264.3	2337.9	2073.7	81.3	28.3	1984–III	4166.4	3102.7	2754.6	141.8	81.7
1973–IV	3289.1	2382.7	2067.4	85.0	29.4	1984–IV	4194.2	3118.5	2784.8	136.3	83.4
1974–I	3259.4	2334.7	2050.8	89.0	29.8	1985–I	4221.8	3123.6	2824.9	125.2	87.2
1974–II	3267.6	2304.5	2059.0	91.2	30.4	1985–II	4254.8	3189.6	2849.7	124.8	90.8
1974–III	3239.1	2315.0	2065.5	97.1	30.9	1985–III	4309.0	3156.5	2893.3	129.8	94.1
1974–IV	3226.4	2313.7	2039.9	86.8	30.5	1985–IV	4333.5	3178.7	2895.3	134.2	97.4
1975–I	3154.0	2282.5	2051.8	75.8	30.0	1986–I	4390.5	3227.5	2922.4	109.2	105.1
1975–II	3190.4	2390.3	2086.9	81.0	29.7	1986–II	4387.7	3281.4	2947.9	106.0	110.7
1975–III	3249.9	2354.4	2114.4	97.8	30.1	1986–III	4412.6	3272.6	2993.7	111.0	112.3
1975–IV	3292.5	2389.4	2137.0	103.4	30.6	1986–IV	4427.1	3266.2	3012.5	119.2	111.0
1976–I	3356.7	2424.5	2179.3	108.4	32.6	1987–I	4460.0	3295.2	3011.5	140.2	108.0
1976–II	3369.2	2434.9	2194.7	109.2	35.0	1987–II	4515.3	3241.7	3046.8	157.9	105.5
1976–III	3381.0	2444.7	2213.0	110.0	36.6	1987–III	4559.3	3285.7	3075.8	169.1	105.1
1976–IV	3416.3	2459.5	2242.0	110.3	38.3	1987–IV	4625.5	3335.8	3074.6	176.0	106.3
1977–I	3466.4	2463.0	2271.3	121.5	39.2	1988–I	4655.3	3380.1	3128.2	195.5	109.6
1977–II	3525.0	2490.3	2280.8	129.7	40.0	1988–II	4704.8	3386.3	3147.8	207.2	113.3
1977–III	3574.4	2541.0	2302.6	135.1	41.4	1988–III	4734.5	3407.5	3170.6	213.4	117.5
1977–IV	3567.2	2556.2	2331.6	134.8	42.4	1988–IV	4779.7	3443.1	3202.9	226.0	121.0
1978–I	3591.8	2587.3	2347.1	137.5	43.5	1989–I	4809.8	3473.9	3200.9	221.3	124.6
1978–II	3707.0	2631.9	2394.0	154.0	44.5	1989–II	4832.4	3450.9	3208.6	206.2	127.1
1978–III	3735.6	2653.2	2404.5	158.0	46.6	1989–III	4845.6	3466.9	3241.1	195.7	129.1
1978–IV	3779.6	2680.9	2421.6	167.8	48.9	1989–IV	4859.7	3493.0	3241.6	203.0	130.7
1979–I	3780.8	2699.2	2437.9	168.2	50.5	1990–1	4880.8	3531.4	3258.8	199.1	132.3
1979–II	3784.3	2697.6	2435.4	174.1	51.8	1990–II	4900.3	3545.3	3258.6	193.7	132.5
1979–III	3807.5	2715.3	2454.7	178.1	52.7	1990–III	4903.3	3547.0	3281.2	196.3	133.8
1979–IV	3814.6	2728.1	2465.4	173.4	54.5	1990–IV	4855.1	3529.5	3251.8	199.0	136.2
1980–I	3830.8	2742.9	2464.6	174.3	57.6	1991–I	4824.0	3514.8	3241.1	189.7	137.8
1980–II	3732.6	2692.0	2414.2	144.5	58.7	1991–II	4840.7	3537.4	3252.4	182.7	136.7
1980–III	3733.5	2722.5	2440.3	151.0	59.3	1991–III	4862.7	3539.9	3271.2	189.6	138.1
1980–IV	3808.5	2777.0	2469.2	154.6	60.5	1991–IV	4868.0	3547.5	3271.1	190.3	138.5

Notes: GDP (Gross Domestic Product), billions of 1987 dollars, p. A-96. PDI (Personal disposable income), billions of 1987 dollars, p. A-112. PCE (Personal consumption expenditure), billions of 1987 dollars, p. A-96. Profits (corporate profits after tax), billions of dollars, p. A-110. Dividends (net corporate dividend payments), billions of dollars, p. A-110.

Source: U.S. Department of Commerce, Bureau of Economic Analysis, *Business Statistics, 1963–1991,* June 1992.

Figure 21.1 is a plot of the data for GDP, PDI, and PCE, and Figure 21.2 presents the other two time series. A visual plot of the data is usually the first step in the analysis of any time series. The first impression that we get from these graphs is that all the time series shown in Figures 21.1 and 21.2 seem to be "trending" upward, albeit with fluctuations. Suppose we wanted to speculate on the shape of these curves over the quarterly period, say, from

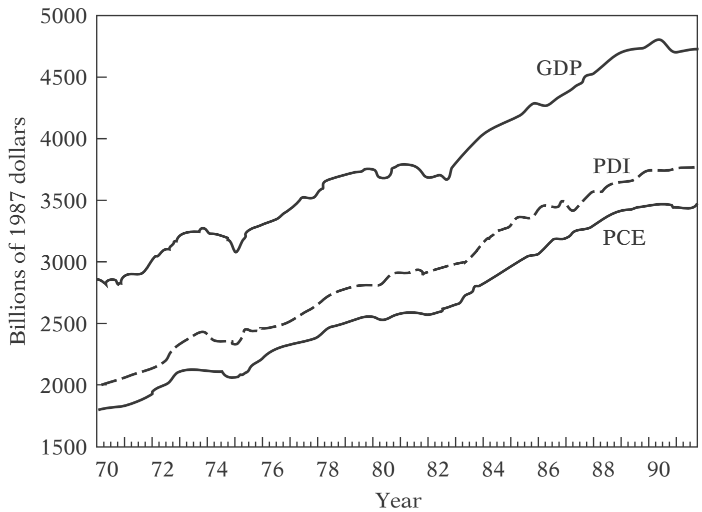

FIGURE 21.1 GDP, PDI, and PCE, United States, 1970–1991 (quarterly).

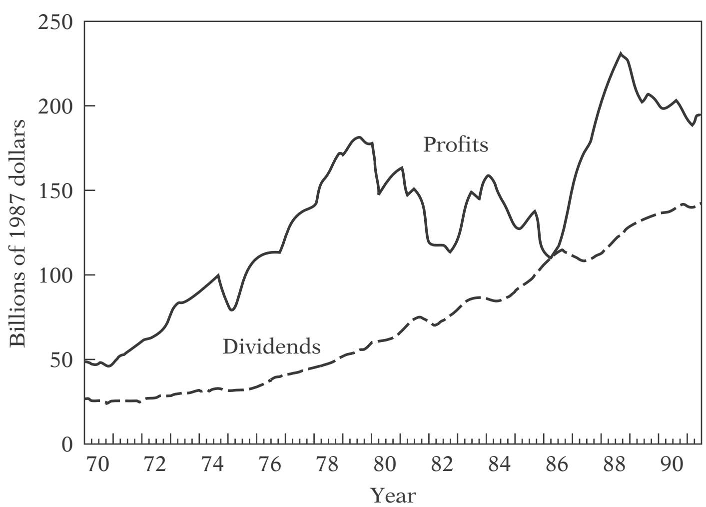

FIGURE 21.2 Profits and dividends, United States, 1970–1991 (quarterly).

1992–I to 1996–IV.[2] Can we simply mentally extend the curves shown in the above figures? Perhaps we can if we know the statistical, or stochastic, mechanism, or the **data generating process (DGP),** that generated these curves? But what is that mechanism? To answer this and related questions, we need to study some "new" vocabulary that has been developed by time series analysts, to which we now turn.

21.2 KEY CONCEPTS[3]

What is this vocabulary? It consists of concepts such as these:

1. Stochastic processes
2. Stationarity processes
3. Purely random processes
4. Nonstationary processes
5. Integrated variables
6. Random walk models
7. Cointegration
8. Deterministic and stochastic trends
9. Unit root tests

In what follows we will discuss each of these concepts. Our discussion will often be heuristic. Wherever possible and helpful, we will provide appropriate examples.

21.3 STOCHASTIC PROCESSES

A random or stochastic process is a collection of random variables ordered in time.[4] If we let Y denote a random variable, and if it is continuous, we denote it as $Y(t)$, but if it is discrete, we denoted it as Y_t. An example of the former is an electrocardiogram, and an example of the latter is GDP, PDI, etc. Since most economic data are collected at discrete points in time, for our purpose we will use the notation Y_t rather than $Y(t)$. If we let Y represent GDP, for our data we have $Y_1, Y_2, Y_3, \ldots, Y_{86}, Y_{87}, Y_{88}$, where the subscript 1 denotes the first observation (i.e., GDP for the first quarter of 1970) and the subscript 88 denotes the last observation (i.e., GDP for the fourth quarter of 1991). *Keep in mind that each of these Y's is a random variable.*

In what sense can we regard GDP as a stochastic process? Consider for instance the GDP of $2872.8 billion for 1970–I. In theory, the GDP figure for

[2] Of course, we have the actual data for this period now and could compare it with the data that is "predicted" on the basis of the earlier period.

[3] The following discussion is based on Maddala et al., op. cit., and Charemza et al., op. cit.

[4] The term "stochastic" comes from the Greek word "stokhos," which means a target or bull's-eye. If you have ever thrown darts on a dart board with the aim of hitting the bull's-eye, how often did you hit the bull's-eye? Out of a hundred darts you may be lucky to hit the bull's-eye only a few times; at other times the darts will be spread randomly around the bull's-eye.

the first quarter of 1970 could have been any number, depending on the economic and political climate then prevailing. The figure of 2872.8 is a particular **realization** of all such possibilities.[5] Therefore, we can say that GDP is a stochastic process and the actual values we observed for the period 1970–I to 1991–IV are a particular realization of that process (i.e., sample). The distinction between the stochastic process and its realization is akin to the distinction between population and sample in cross-sectional data. Just as we use sample data to draw inferences about a population, in time series we use the realization to draw inferences about the underlying stochastic process.

Stationary Stochastic Processes

A type of stochastic process that has received a great deal of attention and scrutiny by time series analysts is the so-called **stationary stochastic process.** Broadly speaking, *a stochastic process is said to be stationary if its mean and variance are constant over time and the value of the covariance between the two time periods depends only on the distance or gap or lag between the two time periods and not the actual time at which the covariance is computed.* In the time series literature, such a stochastic process is known as a **weakly stationary,** or **covariance stationary,** or **second-order stationary,** or **wide sense, stochastic process.** For the purpose of this chapter, and in most practical situations, this type of stationarity often suffices.[6]

To explain weak stationarity, let Y_t be a stochastic time series with these properties:

$$\textit{Mean:} \qquad E(Y_t) = \mu \tag{21.3.1}$$

$$\textit{Variance:} \qquad \operatorname{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \tag{21.3.2}$$

$$\textit{Covariance:} \qquad \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \tag{21.3.3}$$

where γ_k, the covariance (or autocovariance) at lag k, is the covariance between the values of Y_t and Y_{t+k}, that is, between two Y values k periods apart. If $k = 0$, we obtain γ_0, which is simply the variance of $Y\,(=\sigma^2)$; if $k = 1$, γ_1 is the covariance between two adjacent values of Y, the type of covariance we encountered in Chapter 12 (recall the Markov first-order autoregressive scheme).

Suppose we shift the origin of Y from Y_t to Y_{t+m} (say, from the first quarter of 1970 to the first quarter of 1975 for our GDP data). Now if Y_t is to be stationary, the mean, variance, and autocovariances of Y_{t+m} must be the

[5]You can think of the value of $2872.8 billion as the mean value of all possible values of GDP for the first quarter of 1970.

[6]A time series is strictly stationary if *all* the moments of its probability distribution and not just the first two (i.e., mean and variance) are invariant over time. If, however, the stationary process is normal, the weakly stationary stochastic process is also strictly stationary, for the normal stochastic process is fully specified by its two moments, the mean and the variance.

same as those of Y_t. *In short, if a time series is stationary, its mean, variance, and autocovariance (at various lags) remain the same no matter at what point we measure them; that is, they are time invariant*. Such a time series will tend to return to its mean (called **mean reversion**) and fluctuations around this mean (measured by its variance) will have a broadly constant amplitude.[7]

If a time series is not stationary in the sense just defined, it is called a **nonstationary time series** (keep in mind we are talking only about weak stationarity). In other words, a nonstationary time series will have a *time-varying mean or a time-varying variance or both*.

Why are stationary time series so important? Because if a time series is nonstationary, we can study its behavior only for the time period under consideration. Each set of time series data will therefore be for a particular episode. As a consequence, it is not possible to generalize it to other time periods. Therefore, for the purpose of forecasting, such (nonstationary) time series may be of little practical value.

How do we know that a particular time series is stationary? In particular, are the time series shown in Figures 21.1 and 21.2 stationary? We will take this important topic up in Sections 21.8 and 21.9, where we will consider several tests of stationarity. But if we depend on common sense, it would seem that the time series depicted in Figures 21.1 and 21.2 are nonstationary, at least in the mean values. But more on this later.

Before we move on, we mention a special type of stochastic process (or time series), namely, a **purely random,** or **white noise, process.** We call a stochastic process purely random if it has zero mean, constant variance σ^2, and is serially uncorrelated.[8] You may recall that the error term u_t, entering the classical normal linear regression model that we discussed in **Part I** of this book was assumed to be a white noise process, which we denoted as $u_t \sim \text{IIDN}(0, \sigma^2)$; that is, u_t is independently and identically distributed as a normal distribution with zero mean and constant variance.

Nonstationary Stochastic Processes

Although our interest is in stationary time series, one often encounters nonstationary time series, the classic example being the **random walk model** (RWM).[9] It is often said that asset prices, such as stock prices or exchange rates, follow a random walk; that is, they are nonstationary. We distinguish two types of random walks: (1) random walk without drift (i.e., no constant or intercept term) and (2) random walk with drift (i.e., a constant term is present).

---

[7]This point has been made by Keith Cuthbertson, Stephen G. Hall, and Mark P. Taylor, *Applied Econometric Techniques,* The University of Michigan Press, 1995, p. 130.

[8]If it is also independent, such a process is called **strictly white noise.**

[9]The term random walk is often compared with a drunkard's walk. Leaving a bar, the drunkard moves a random distance u_t at time t, and, continuing to walk indefinitely, will eventually drift farther and farther away from the bar. The same is said about stock prices. Today's stock price is equal to yesterday's stock price plus a random shock.

Random Walk without Drift. Suppose u_t is a white noise error term with mean 0 and variance σ^2. Then the series Y_t is said to be a random walk if

$$Y_t = Y_{t-1} + u_t \tag{21.3.4}$$

In the random walk model, as (21.3.4) shows, the value of Y at time t is equal to its value at time $(t-1)$ plus a random shock; thus it is an AR(1) model in the language of Chapters 12 and 17. We can think of (21.3.4) as a regression of Y at time t on its value lagged one period. Believers in the **efficient capital market hypothesis** argue that stock prices are essentially random and therefore there is no scope for profitable speculation in the stock market: If one could predict tomorrow's price on the basis of today's price, we would all be millionaires.

Now from (21.3.4) we can write

$$Y_1 = Y_0 + u_1$$
$$Y_2 = Y_1 + u_2 = Y_0 + u_1 + u_2$$
$$Y_3 = Y_2 + u_3 = Y_0 + u_1 + u_2 + u_3$$

In general, if the process started at some time 0 with a value of Y_0, we have

$$Y_t = Y_0 + \sum u_t \tag{21.3.5}$$

Therefore,

$$E(Y_t) = E\left(Y_0 + \sum u_t\right) = Y_0 \quad \text{(why?)} \tag{21.3.6}$$

In like fashion, it can be shown that

$$\text{var}(Y_t) = t\sigma^2 \tag{21.3.7}$$

As the preceding expression shows, the mean of Y is equal to its initial, or starting, value, which is constant, but as t increases, its variance increases indefinitely, thus violating a condition of stationarity. In short, the RWM without drift is a nonstationary stochastic process. In practice Y_0 is often set at zero, in which case $E(Y_t) = 0$.

An interesting feature of RWM is the *persistence of random shocks* (i.e., random errors), which is clear from (21.3.5): Y_t is the sum of initial Y_0 plus the sum of random shocks. As a result, the impact of a particular shock does not die away. For example, if $u_2 = 2$ rather than $u_2 = 0$, then all Y_t's from Y_2 onward will be 2 units higher and the effect of this shock never dies out. That is why random walk is said to have an *infinite memory*. As Kerry Patterson notes, random walk remembers the shock forever[10]; that is, it has infinite memory.

[10] Kerry Patterson, op cit., Chap. 6.

Interestingly, if you write (21.3.4) as

$$(Y_t - Y_{t-1}) = \Delta Y_t = u_t \tag{21.3.8}$$

where Δ is the first difference operator that we discussed in Chapter 12. It is easy to show that, while Y_t is nonstationary, its first difference is stationary. In other words, the first differences of a random walk time series are stationary. But we will have more to say about this later.

Random Walk with Drift. Let us modify (21.3.4) as follows:

$$Y_t = \delta + Y_{t-1} + u_t \tag{21.3.9}$$

where δ is known as the **drift parameter.** The name drift comes from the fact that if we write the preceding equation as

$$Y_t - Y_{t-1} = \Delta Y_t = \delta + u_t \tag{21.3.10}$$

it shows that Y_t drifts upward or downward, depending on δ being positive or negative. Note that model (21.3.9) is also an AR(1) model.

Following the procedure discussed for random walk without drift, it can be shown that for the random walk with drift model (21.3.9),

$$E(Y_t) = Y_0 + t \cdot \delta \tag{21.3.11}$$

$$\operatorname{var}(Y_t) = t\sigma^2 \tag{21.3.12}$$

As you can see, for RWM with drift the mean as well as the variance increases over time, again violating the conditions of (weak) stationarity. In short, RWM, with or without drift, is a nonstationary stochastic process.

To give a glimpse of the random walk with and without drift, we conducted two simulations as follows:

$$Y_t = Y_0 + u_t \tag{21.3.13}$$

where u_t are white noise error terms such that each $u_t \sim N(0, 1)$; that is, each u_t follows the standard normal distribution. From a random number generator, we obtained 500 values of u and generated Y_t as shown in (21.3.13). We assumed $Y_0 = 0$. Thus, (21.3.13) is an RWM without drift.

Now consider

$$Y_t = \delta + Y_0 + u_t \tag{21.3.14}$$

which is RWM with drift. We assumed u_t and Y_0 as in (21.3.13) and assumed that $\delta = 2$.

The graphs of models (21.3.13) and (21.3.14), respectively, are in Figures 21.3 and 21.4. The reader can compare these two diagrams in light of our discussion of the RWM with and without drift.

The random walk model is an example of what is known in the literature as a **unit root process.** Since this term has gained tremendous currency in the time series literature, we next explain what a unit root process is.

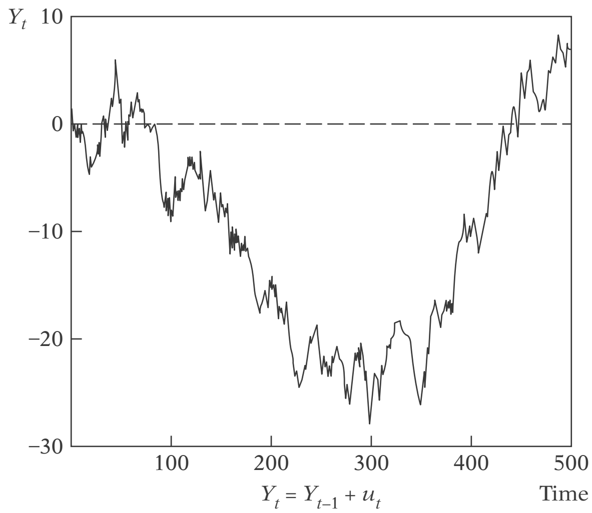

FIGURE 21.3 A random walk without drift.

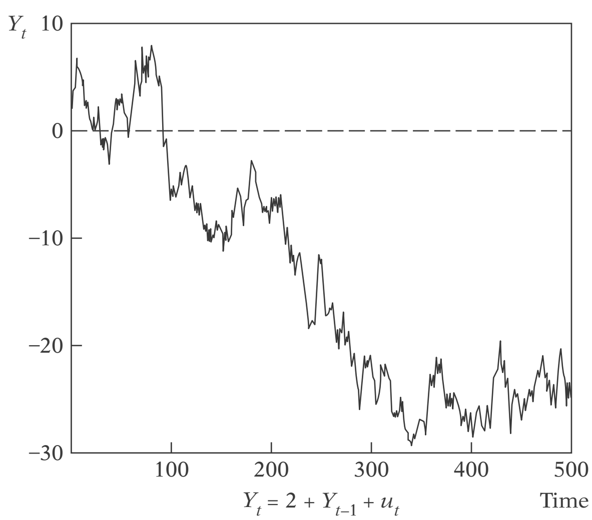

FIGURE 21.4 A random walk with drift.

21.4 UNIT ROOT STOCHASTIC PROCESS

Let us write the RWM (21.3.4) as:

$$Y_t = \rho Y_{t-1} + u_t \qquad -1 \leq \rho \leq 1 \tag{21.4.1}$$

This model resembles the Markov first-order autoregressive model that we discussed in the chapter on autocorrelation. If $\rho = 1$, (21.4.1) becomes a RWM (without drift). If ρ is in fact 1, we face what is known as the **unit root problem,** that is, a situation of nonstationarity; we already know that in this case the variance of Y_t is not stationary. The name unit root is due to the fact that $\rho = 1$.[11] *Thus the terms nonstationarity, random walk, and unit root can be treated as synonymous.*

If, however, $|\rho| \leq 1$, that is if the absolute value of ρ is less than one, then it can be shown that the time series Y_t is stationary in the sense we have defined it.[12]

In practice, then, it is important to find out if a time series possesses a unit root.[13] In Section 21.9 we will discuss several tests of unit root, that is, several tests of stationarity. In that section we will also determine whether the time series depicted in Figures 21.1 and 21.2 are stationary. Perhaps the reader might suspect that they are not. But we shall see.

21.5 TREND STATIONARY (TS) AND DIFFERENCE STATIONARY (DS) STOCHASTIC PROCESSES

The distinction between stationary and nonstationary stochastic processes (or time series) has a crucial bearing on whether the trend (the slow long-run evolution of the time series under consideration) observed in the constructed time series in Figures 21.3 and 21.4 or in the actual economic time series of Figures 21.1 and 21.2 is **deterministic** or **stochastic.** Broadly speaking, if the trend in a time series is completely predictable and not variable, we call it a deterministic trend, whereas if it is not predictable, we call it a stochastic trend. To make the definition more formal, consider the following model of the time series Y_t.

$$Y_t = \beta_1 + \beta_2 t + \beta_3 Y_{t-1} + u_t \tag{21.5.1}$$

---

[11]A technical point: If $\rho = 1$, we can write (21.4.1) as $Y_t - Y_{t-1} = u_t$. Now using the **lag operator** L so that $LY_t = Y_{t-1}$, $L^2Y_t = Y_{t-2}$, and so on, we can write (21.4.1) as $(1 - L)Y_t = u_t$. The term unit root refers to the root of the polynomial in the lag operator. If you set $(1 - L) = 0$, we obtain, $L = 1$, hence the name unit root.

[12]If in (21.4.1) it is assumed that the initial value of $Y\ (= Y_0)$ is zero, $|\rho| \leq 1$, and u_t is white noise and distributed normally with zero mean and unit variance, then it follows that $E(Y_t) = 0$ and $\text{var}\,(Y_t) = 1/(1 - \rho^2)$. Since both these are constants, by the definition of weak stationarity, Y_t is stationary. On the other hand, as we saw before, if $\rho = 1$, Y_t is a random walk or nonstationary.

[13]A time series may contain more than one unit root. But we will discuss this situation later in the chapter.

where u_t is a white noise error term and where t is time measured chronologically. Now we have the following possibilities:

Pure random walk: If in (21.5.1) $\beta_1 = 0, \beta_2 = 0, \beta_3 = 1,$ we get

$$Y_t = Y_{t-1} + u_t \tag{21.5.2}$$

which is nothing but a RWM without drift and is therefore nonstationary. But note that, if we write (21.5.2) as

$$\Delta Y_t = (Y_t - Y_{t-1}) = u_t \tag{21.3.8}$$

it becomes stationary, as noted before. Hence, a RWM without drift is a **difference stationary process (DSP).**

Random walk with drift: If in (21.5.1) $\beta_1 \neq 0, \beta_2 = 0, \beta_3 = 1,$ we get

$$Y_t = \beta_1 + Y_{t-1} + u_t \tag{21.5.3}$$

which is a random walk with drift and is therefore nonstationary. If we write it as

$$(Y_t - Y_{t-1}) = \Delta Y_t = \beta_1 + u_t \tag{21.5.3a}$$

this means Y_t will exhibit a positive ($\beta_1 > 0$) or negative ($\beta_1 < 0$) trend (see Figure 21.4). Such a trend is called a **stochastic trend.** Equation (21.5.3*a*) is a DSP process because the nonstationarity in Y_t can be eliminated by taking first differences of the time series.

Deterministic trend: If in (21.5.1), $\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 = 0,$ we obtain

$$Y_t = \beta_1 + \beta_2 t + u_t \tag{21.5.4}$$

which is called a **trend stationary process (TSP).** Although the mean of Y_t is $\beta_1 + \beta_2 t$, which is not constant, its variance ($= \sigma^2$) is. Once the values of β_1 and β_2 are known, the mean can be forecast perfectly. Therefore, if we subtract the mean of Y_t from Y_t, the resulting series will be stationary, hence the name **trend stationary.** This procedure of removing the (deterministic) trend is called **detrending.**

Random walk with drift and deterministic trend: If in (21.5.1), $\beta_1 \neq 0,$ $\beta_2 \neq 0, \beta_3 = 1,$ we obtain:

$$Y_t = \beta_1 + \beta_2 t + Y_{t-1} + u_t \tag{21.5.5}$$

we have a random walk with drift and a deterministic trend, which can be seen if we write this equation as

$$\Delta Y_t = \beta_1 + \beta_2 t + u_t \tag{21.5.5a}$$

which means that Y_t is nonstationary.

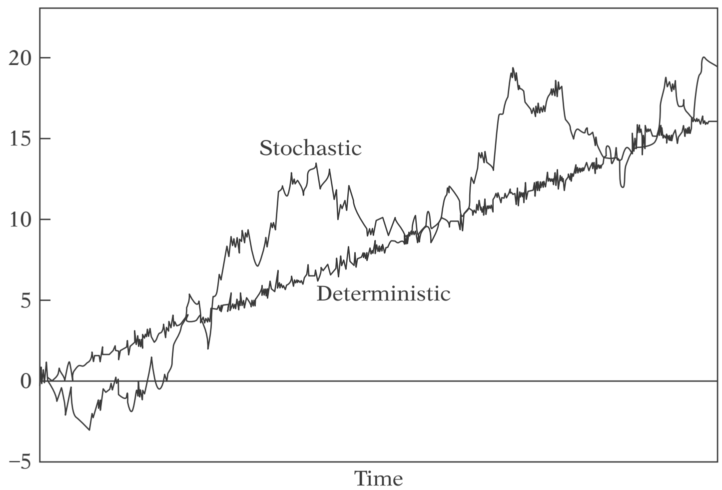

FIGURE 21.5 Deterministic versus stochastic trend.
Source: Charemza et al., op. cit., p. 91.

Deterministic trend with stationary AR(1) component: If in (21.5.1) $\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 < 1$, then we get

$$Y_t = \beta_1 + \beta_2 t + \beta_3 Y_{t-1} + u_t \quad \textbf{(21.5.6)}$$

which is stationary around the deterministic trend.

To see the difference between stochastic and deterministic trends, consider Figure 21.5.[14] The series named stochastic in this figure is generated by an RWM: $Y_t = 0.5 + Y_{t-1} + u_t$, where 500 values of u_t were generated from a standard normal distribution and where the initial value of Y was set at 1. The series named deterministic is generated as follows: $Y_t = 0.5t + u_t$, where u_t were generated as above and where t is time measured chronologically.

As you can see from Figure 21.5, in the case of the deterministic trend, the deviations from the trend line (which represents nonstationary mean) are purely random and they die out quickly; they do not contribute to the long-run development of the time series, which is determined by the trend component $0.5t$. In the case of the stochastic trend, on the other hand, the random component u_t affects the long-run course of the series Y_t.

21.6 INTEGRATED STOCHASTIC PROCESSES

The random walk model is but a specific case of a more general class of stochastic processes known as **integrated processes.** Recall that the RWM without drift is nonstationary, but its first difference, as shown in (21.3.8), is

[14]The following discussion is based on Wojciech W. Charemza et al., op. cit., pp. 89–91.