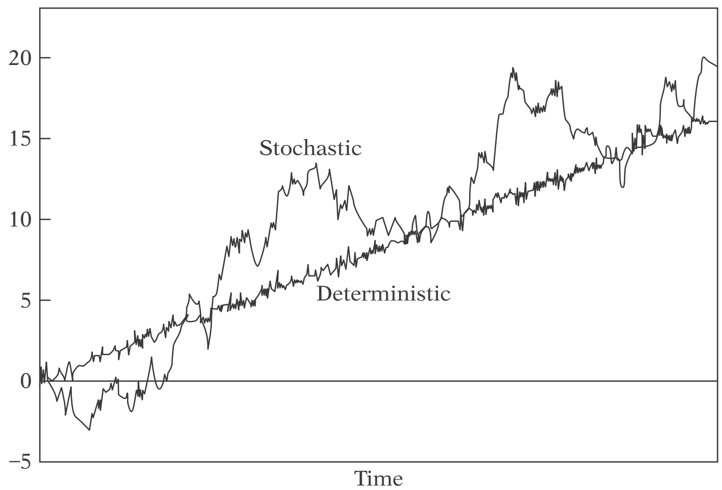


FIGURE 21.5 Deterministic versus stochastic trend.
Source: Charemza et al., op. cit., p. 91.

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

$$Y_t = \beta_1 + \beta_2 t + \beta_3 Y_{t-1} + u_t \tag{21.5.6}$$

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.

stationary. Therefore, we call the RWM without drift **integrated of order 1,** denoted as $I(1)$. Similarly, if a time series has to be differenced twice (i.e., take the first difference of the first differences) to make it stationary, we call such a time series **integrated of order 2.**[15] In general, if a (nonstationary) time series has to be differenced d times to make it stationary, that time series is said to be **integrated of order *d*.** A time series Y_t integrated of order d is denoted as $Y_t \sim I(d)$. If a time series Y_t is stationary to begin with (i.e., it does not require any differencing), it is said to be integrated of order zero, denoted by $Y_t \sim I(0)$. Thus, we will use the terms "stationary time series" and "time series integrated of order zero" to mean the same thing.

Most economic time series are generally $I(1)$; that is, they generally become stationary only after taking their first differences. Are the time series shown in Figures 21.1 and 21.2 $I(1)$ or of higher order? We will examine them in Sections 21.8 and 21.9.

Properties of Integrated Series

The following properties of integrated time series may be noted: Let X_t, Y_t, and Z_t be three time series.

1. If $X_t \sim I(0)$ and $Y_t \sim I(1)$, then $Z_t = (X_t + Y_t) = I(1)$; that is, a linear combination or sum of stationary and nonstationary time series is nonstationary.

2. If $X_t \sim I(d)$, then $Z_t = (a + bX_t) = I(d)$, where a and b are constants. That is, a linear combination of an $I(d)$ series is also $I(d)$. Thus, if $X_t \sim I(0)$, then $Z_t = (a + bX_t) \sim I(0)$.

3. If $X_t \sim I(d_1)$ and $Y_t \sim I(d_2)$, then $Z_t = (aX_t + bY_t) \sim I(d_2)$, where $d_1 < d_2$.

4. If $X_t \sim I(d)$ and $Y_t \sim I(d)$, then $Z_t = (aX_t + bY_t) \sim I(d^*)$; d^* is generally equal to d, but in some cases $d^* < d$ (see the topic of cointegration in Section 21.11).

As you can see from the preceding statements, one has to pay careful attention in combining two or more time series that are integrated of different order.

To see why this is important, consider the two-variable regression model discussed in Chapter 3, namely, $Y_t = \beta_1 + \beta_2 X_t + u_t$. Under the classical OLS assumptions, we know that

$$\hat{\beta}_2 = \frac{\sum x_t y_t}{\sum x_t^2} \qquad \textbf{(21.6.1)}$$

where the small letters, as usual, indicate deviation from mean values. Suppose Y_t is $I(0)$, but X_t is $I(1)$; that is, the former is stationary and the latter is

[15]For example if Y_t is $I(2)$, then $\Delta\Delta Y_t = \Delta(Y_t - Y_{t-1}) = \Delta Y_t - \Delta Y_{t-1} = Y_t - 2Y_{t-1} + Y_{t-2}$ will become stationary. But note that $\Delta\Delta Y_t = \Delta^2 Y_t \neq Y_t - Y_{t-2}$.

not. Since X_t is nonstationary, its variance will increase indefinitely, thus dominating the numerator term in (21.6.1) with the result that $\hat{\beta}_2$ will converge to zero asymptotically (i.e., in large samples) and it will not even have an asymptotic distribution.[16]

21.7 THE PHENOMENON OF SPURIOUS REGRESSION

To see why stationary time series are so important, consider the following two random walk models:

$$Y_t = Y_{t-1} + u_t \qquad \textbf{(21.7.1)}$$

$$X_t = X_{t-1} + v_t \qquad \textbf{(21.7.2)}$$

where we generated 500 observations of u_t from $u_t \sim N(0, 1)$ and 500 observations of v_t from $v_t \sim N(0, 1)$ and assumed that the initial values of both Y and X were zero. We also assumed that u_t and v_t are serially uncorrelated as well as mutually uncorrelated. As you know by now, both these time series are nonstationary; that is, they are $I(1)$ or exhibit stochastic trends.

Suppose we regress Y_t on X_t. Since Y_t and X_t are uncorrelated $I(1)$ processes, the R^2 from the regression of Y on X should tend to zero; that is, there should not be any relationship between the two variables. But wait till you see the regression results:

Variable	Coefficient	Std. error	*t* statistic
C	-13.2556	0.6203	-21.36856
X	0.3376	0.0443	7.61223
	R^2 = 0.1044	d = 0.0121	

As you can see, the coefficient of X is highly statistically significant, and, although the R^2 value is low, it is statistically significantly different from zero. From these results, you may be tempted to conclude that there is a significant statistical relationship between Y and X, whereas a priori there should be none. This is in a nutshell the **phenomenon of spurious** or **nonsense regression,** first discovered by Yule.[17] Yule showed that (spurious) correlation could persist in nonstationary time series even if the sample is very large. That there is something wrong in the preceding regression is suggested by the extremely low Durbin–Watson d value, which suggests very

[16]This point is due to Maddala et al., op. cit., p. 26.

[17]Yule, G. U., "Why Do We Sometimes Get Nonsense Correlations Between Time Series? A Study in Sampling and the Nature of Time Series," *Journal of the Royal Statistical Society,* vol. 89, 1926, pp. 1–64. For extensive Monte Carlo simulations on spurious regression see C. W. J. Granger and P. Newbold, "Spurious Regressions in Econometrics," *Journal of Econometrics,* vol. 2, 1974, pp. 111–120.

strong first-order autocorrelation. According to Granger and Newbold, *an $R^2 > d$ is a good rule of thumb to suspect that the estimated regression is spurious,* as in the example above.

That the regression results presented above are meaningless can be easily seen from regressing the first differences of Y_t $(= \Delta Y_t)$ on the first differences of X_t $(= \Delta X_t)$; remember that although Y_t and X_t are nonstationary, their first differences are stationary. In such a regression you will find that R^2 is practically zero, as it should be, and the Durbin–Watson d is about 2. In Exercise 21.24 you are asked to run this regression and verify the statement just made.

Although dramatic, this example is a strong reminder that one should be extremely wary of conducting regression analysis based on time series that exhibit stochastic trends. And one should therefore be extremely cautious in reading too much in the regression results based on $I(1)$ variables. For an example, see exercise 21.26. To some extent, this is true of time series subject to deterministic trends, an example of which is given in exercise 21.25.

21.8 TESTS OF STATIONARITY

By now the reader probably has a good idea about the nature of stationary stochastic processes and their importance. In practice we face two important questions: (1) How do we find out if a given time series is stationary? (2) If we find that a given time series is not stationary, is there a way that it can be made stationary? We take up the first question in this section and discuss the second question in Section 21.10.

Before we proceed, keep in mind that we are primarily concerned with weak, or covariance, stationarity.

Although there are several tests of stationarity, we discuss only those that are prominently discussed in the literature. In this section we discuss two tests: (1) graphical analysis and (2) the correlogram test. Because of the importance attached to it in the recent past, we discuss the *unit root test* in the next section. We illustrate these tests with appropriate examples.

1. Graphical Analysis

As noted earlier, before one pursues formal tests, it is always advisable to plot the time series under study, as we have done in Figures 21.1 and 21.2 for the data given in Table 21.1. Such a plot gives an initial clue about the likely nature of the time series. Take, for instance, the GDP time series shown in Figure 21.1. You will see that over the period of study GDP has been increasing, that is, showing an upward trend, suggesting perhaps that the mean of the GDP has been changing. This perhaps suggests that the GDP series is not stationary. This is also more or less true of the other U.S. economic time series shown in Figure 21.2. Such an intuitive feel is the starting point of more formal tests of stationarity.

2. Autocorrelation Function (ACF) and Correlogram

One simple test of stationarity is based on the so-called **autocorrelation function (ACF).** The ACF at lag k, denoted by ρ_k, is defined as

$$\begin{aligned}\rho_k &= \frac{\gamma_k}{\gamma_0} \\ &= \frac{\text{covariance at lag } k}{\text{variance}}\end{aligned} \tag{21.8.1}$$

where covariance at lag k and variance are as defined before. Note that if $k = 0$, $\rho_0 = 1$ (why?)

Since both covariance and variance are measured in the same units of measurement, ρ_k is a *unitless,* or *pure, number.* It lies between −1 and +1, as any correlation coefficient does. If we plot ρ_k against k, the graph we obtain is known as the **population correlogram.**

Since in practice we only have a realization (i.e., sample) of a stochastic process, we can only compute the **sample autocorrelation function (SAFC),** $\hat{\rho}_k$. To compute this, we must first compute the **sample covariance** at lag k, $\hat{\gamma}_k$, and the **sample variance,** $\hat{\gamma}_0$, which are defined as[18]

$$\hat{\gamma}_k = \frac{\sum (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{n} \tag{21.8.2}$$

$$\hat{\gamma}_0 = \frac{\sum (Y_t - \bar{Y})^2}{n} \tag{21.8.3}$$

where n is the sample size and $\bar{Y}$ is the sample mean.

Therefore, the sample autocorrelation function at lag k is

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \tag{21.8.4}$$

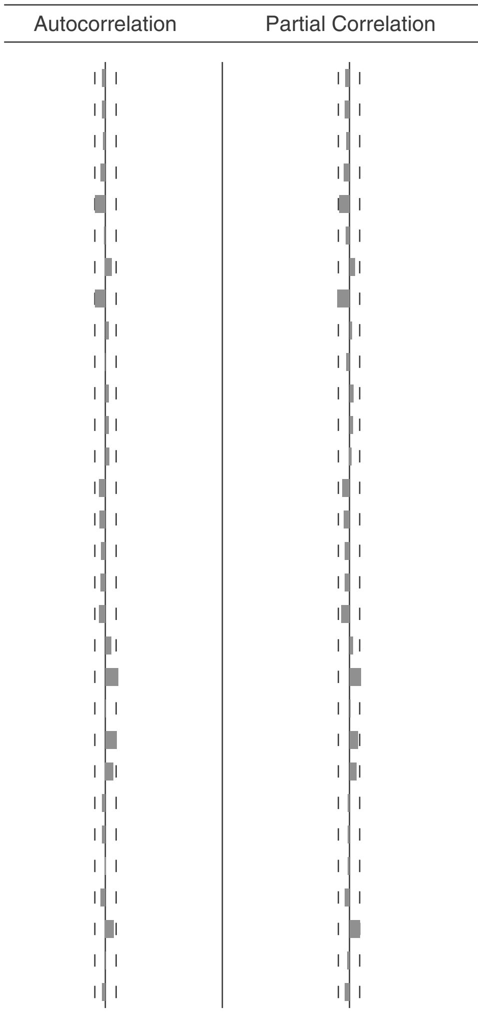
which is simply the ratio of sample covariance (at lag k) to sample variance. A plot of $\hat{\rho}_k$ against k is known as the **sample correlogram.**

How does a sample correlogram enable us to find out if a particular time series is stationary? For this purpose, let us first present the sample correlograms of a purely white noise random process and of a random walk process. Return to the driftless RWM (21.3.13). There we generated a sample of 500 error terms, the *u's,* from the standard normal distribution. The correlogram of these 500 purely random error terms is as shown in Figure 21.6; we have shown this correlogram up to 30 lags. We will comment shortly on how one chooses the lag length.

For the time being, just look at the column labeled AC, which is the sample autocorrelation function, and the first diagram on the left, labeled

[18]Strictly speaking, we should divide the sample covariance at lag k by $(n - k)$ and the sample variance by $(n - 1)$ rather than by n (why?) where n is the sample size.

Sample: 2 500
Included observations: 499



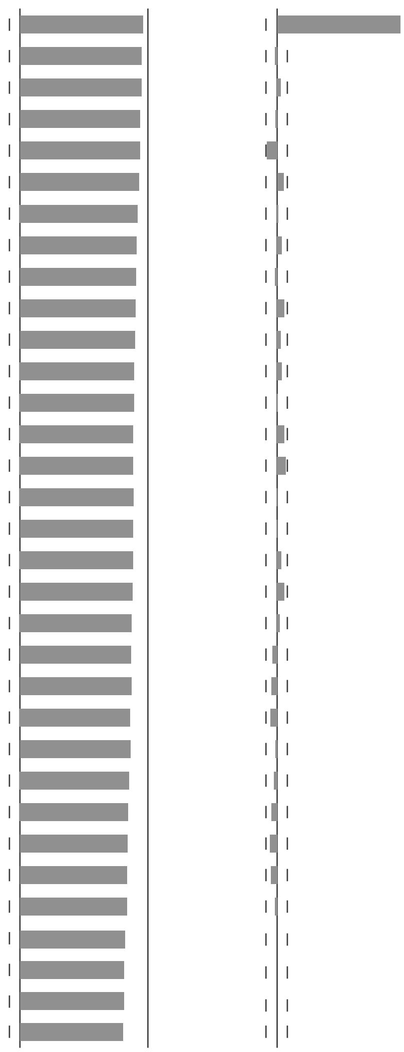
	AC	PAC	Q-Stat	Prob
1	–0.022	–0.022	0.2335	0.629
2	–0.019	–0.020	0.4247	0.809
3	–0.009	–0.010	0.4640	0.927
4	–0.031	–0.031	0.9372	0.919
5	–0.070	–0.072	3.4186	0.636
6	–0.008	–0.013	3.4493	0.751
7	0.048	0.045	4.6411	0.704
8	–0.069	–0.070	7.0385	0.532
9	0.022	0.017	7.2956	0.606
10	–0.004	–0.011	7.3059	0.696
11	0.024	0.025	7.6102	0.748
12	0.024	0.027	7.8993	0.793
13	0.026	0.021	8.2502	0.827
14	–0.047	–0.046	9.3726	0.806
15	–0.037	–0.030	10.074	0.815
16	–0.026	–0.031	10.429	0.843
17	–0.029	–0.024	10.865	0.863
18	–0.043	–0.050	11.807	0.857
19	0.038	0.028	12.575	0.860
20	0.099	0.093	17.739	0.605
21	0.001	0.007	17.739	0.665
22	0.065	0.060	19.923	0.588
23	0.053	0.055	21.404	0.556
24	–0.017	–0.004	21.553	0.606
25	–0.024	–0.005	21.850	0.644
26	–0.008	–0.008	21.885	0.695
27	–0.036	–0.027	22.587	0.707
28	0.053	0.072	24.068	0.678
29	–0.004	–0.011	24.077	0.725
30	–0.026	–0.025	24.445	0.752

FIGURE 21.6 Correlogram of white noise error term *u*. AC = autocorrelation, PAC = partial autocorrelation (see Chapter 22), Q-Stat = *Q* statistic, Prob = probability.

autocorrelation. The solid vertical line in this diagram represents the zero axis; observations above the line are positive values and those below the line are negative values. As is very clear from this diagram, for a purely white noise process the autocorrelations at various lags hover around zero. *This is the picture of a correlogram of a stationary time series.* Thus, if the correlogram of an actual (economic) time series resembles the correlogram of a white noise time series, we can say that time series is probably stationary.

Now look at the correlogram of a random walk series, as generated, say, by (21.3.13). The picture is as shown in Figure 21.7. The most striking feature of this correlogram is that the autocorrelation coefficients at various lags are very high even up to a lag of 33 quarters. As a matter of fact, if we consider lags of up to 60 quarters, the autocorrelation coefficients are quite high; the

Sample: 2 500
Included observations: 499



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.992	0.992	493.86	0.000
		2	0.984	0.000	980.68	0.000
		3	0.976	0.030	1461.1	0.000
		4	0.969	0.005	1935.1	0.000
		5	0.961	–0.059	2402.0	0.000
		6	0.953	0.050	2862.7	0.000
		7	0.946	0.004	3317.3	0.000
		8	0.939	0.040	3766.4	0.000
		9	0.932	–0.009	4210.1	0.000
		10	0.927	0.055	4649.1	0.000
		11	0.921	0.018	5083.9	0.000
		12	0.916	0.039	5514.9	0.000
		13	0.912	0.002	5942.4	0.000
		14	0.908	0.056	6367.0	0.000
		15	0.905	0.061	6789.8	0.000
		16	0.902	0.000	7210.6	0.000
		17	0.899	0.006	7629.4	0.000
		18	0.896	0.030	8046.7	0.000
		19	0.894	0.053	8463.1	0.000
		20	0.892	0.013	8878.7	0.000
		21	0.890	–0.041	9292.6	0.000
		22	0.886	–0.040	9704.1	0.000
		23	0.882	–0.044	10113.	0.000
		24	0.878	–0.012	10518.	0.000
		25	0.873	–0.023	10920.	0.000
		26	0.867	–0.041	11317.	0.000
		27	0.860	–0.055	11709.	0.000
		28	0.853	–0.045	12095.	0.000
		29	0.846	–0.010	12476.	0.000
		30	0.839	0.008	12851.	0.000
		31	0.832	–0.006	13221.	0.000
		32	0.825	0.003	13586.	0.000
		33	0.819	–0.006	13946.	0.000

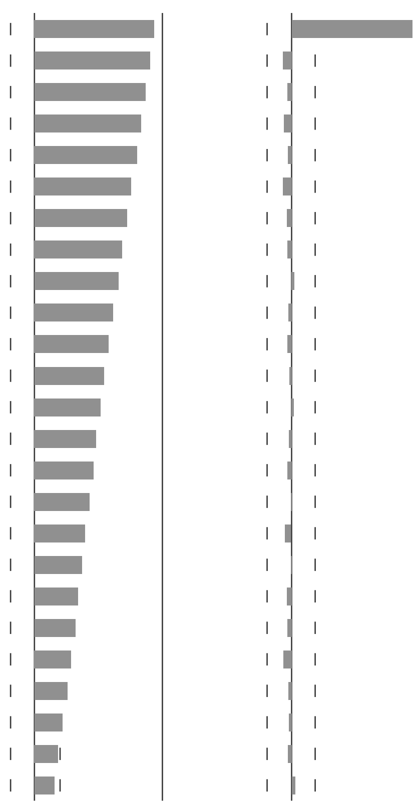
FIGURE 21.7 Correlogram of a random walk time series. See Figure 21.6 for definitions.

coefficient is about 0.7 at lag 60. Figure 21.7 is the typical correlogram of a nonstationary time series: The autocorrelation coefficient starts at a very high value and declines very slowly toward zero as the lag lengthens.

Now let us take a concrete economic example. Let us examine the correlogram of the GDP time series given in Table 21.1. The correlogram up to 25 lags is shown in Figure 21.8. The GDP correlogram up to 25 lags also shows a pattern similar to the correlogram of the random walk model in Figure 21.7. The autocorrelation coefficient starts at a very high value at lag 1 (0.969) and declines very slowly. Thus it seems that the GDP time series is nonstationary. If you plot the correlograms of the other U.S. economic time series shown in Figures 21.1 and 21.2, you will also see a similar pattern, leading to the conclusion that all these time series are nonstationary; they may be nonstationary in mean or variance or both.

Sample: 1970–1 1991–4
Included observations: 88

Autocorrelation | Partial Correlation



	AC	PAC	Q-Stat	Prob
1	0.969	0.969	85.462	0.000
2	0.935	–0.058	166.02	0.000
3	0.901	–0.020	241.72	0.000
4	0.866	–0.045	312.39	0.000
5	0.830	–0.024	378.10	0.000
6	0.791	–0.062	438.57	0.000
7	0.752	–0.029	493.85	0.000
8	0.713	–0.024	544.11	0.000
9	0.675	0.009	589.77	0.000
10	0.638	–0.010	631.12	0.000
11	0.601	–0.020	668.33	0.000
12	0.565	–0.012	701.65	0.000
13	0.532	0.020	731.56	0.000
14	0.500	–0.012	758.29	0.000
15	0.468	–0.021	782.02	0.000
16	0.437	–0.001	803.03	0.000
17	0.405	–0.041	821.35	0.000
18	0.375	–0.005	837.24	0.000
19	0.344	–0.038	850.79	0.000
20	0.313	–0.017	862.17	0.000
21	0.279	–0.066	871.39	0.000
22	0.246	–0.019	878.65	0.000
23	0.214	–0.008	884.22	0.000
24	0.182	–0.018	888.31	0.000
25	0.153	0.017	891.25	0.000

FIGURE 21.8 Correlogram of U.S. GDP, 1970–I to 1991–IV. See Figure 21.6 for definitions.

Two practical questions may be posed here. First, how do we choose the lag length to compute the ACF? Second, how do you decide whether a correlation coefficient at a certain lag is statistically significant? The answer follows.

The Choice of Lag Length. This is basically an empirical question. A rule of thumb is to compute ACF up to one-third to one-quarter the length of the time series. Since for our economic data we have 88 quarterly observations, by this rule lags of 22 to 29 quarters will do. The best practical advice is to start with sufficiently large lags and then reduce them by some statistical criterion, such as the *Akaike* or *Schwarz information criterion* that we discussed in Chapter 13. Alternatively, one can use the following statistical tests.

Statistical Significance of Autocorrelation Coefficients

Consider, for instance, the correlogram of the GDP time series given in Figure 21.8. How do we decide whether the correlation coefficient of 0.638 at lag 10 (quarters) is statistically significant? The statistical significance of any $\hat{\rho}_k$ can be judged by its standard error. Bartlett has shown that if a time series is purely random, that is, it exhibits white noise (see Figure 21.6), the sample autocorrelation coefficients $\hat{\rho}_k$ are *approximately*[19]

$$\hat{\rho}_k \sim N(0, 1/n) \tag{21.8.5}$$

that is, in large samples the sample autocorrelation coefficients are normally distributed with zero mean and variance equal to one over the sample size. Since we have 88 observations, the variance is $1/88 = 0.01136$ and the standard error is $\sqrt{0.01136} = 0.1066$. Then following the properties of the standard normal distribution, the 95% confidence interval for any (population) ρ_k is:

$$\hat{\rho}_k \pm 1.96(0.1066) \tag{21.8.6}$$

In other words,

$$\text{Prob}\,(\hat{\rho}_k - 0.2089 \le \rho_k \le \hat{\rho}_k + 0.2089) = 0.95 \tag{21.8.7}$$

If the preceding interval includes the value of zero, we do not reject the hypothesis that the true ρ_k is zero, but if this interval does not include 0, we reject the hypothesis that the true ρ_k is zero. Applying this to the estimated value of $\hat{\rho}_{10} = 0.638$, the reader can verify that the 95% confidence interval for true ρ_{10} is (0.638 ± 0.2089) or $(0.4291, 0.8469)$.[20] Obviously, this inter-

[19] M. S. Bartlett, "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series," *Journal of the Royal Statistical Society,* Series B, vol. 27, 1946, pp. 27–41.

val does not include the value of zero, suggesting that we are 95% confident that the true ρ_{10} is significantly different from zero.[21] As you can check, even at lag 20 the estimated ρ_{20} is statistically significant at the 5% level.

Instead of testing the statistical significance of any individual autocorrelation coefficient, we can test the *joint hypothesis* that all the ρ_k up to certain lags are simultaneously equal to zero. This can be done by using the ***Q* statistic** developed by Box and Pierce, which is defined as[22]

$$Q = n\sum_{k=1}^{m}\hat{\rho}_k^2 \tag{21.8.8}$$

where n = sample size and m = lag length. The Q statistic is often used as a test of whether a time series is white noise. In large samples, it is *approximately* distributed as the chi-square distribution with m df. In an application, if the computed Q exceeds the critical Q value from the chi-square distribution at the chosen level of significance, one can reject the null hypothesis that all the (true) ρ_k are zero; at least some of them must be nonzero.

A variant of the Box–Pierce Q statistic is the **Ljung–Box (LB) statistic,** which is defined as[23]

$$\text{LB} = n(n+2)\sum_{k=1}^{m}\left(\frac{\hat{\rho}_k^2}{n-k}\right) \sim \chi^2 m \tag{21.8.9}$$

Although in large samples both Q and LB statistics follow the chi-square distribution with m df, the LB statistic has been found to have better (more powerful, in the statistical sense) small-sample properties than the Q statistic.[24]

Returning to the GDP example given in Figure 21.8, the value of the LB statistic up to lag 25 is about 891.25. The probability of obtaining such an LB value under the null hypothesis that the sum of 25 squared estimated autocorrelation coefficients is zero is practically zero, as the last column of that figures shows. Therefore, the conclusion is that the GDP time series is nonstationary, therefore reinforcing our hunch from Figure 21.1 that the GDP series may be nonstationary. In exercise 21.16 you are asked to confirm that the other four U.S. economic time series are also nonstationary.

[20]Our sample size of 88 observations, although not very large, is reasonably large to use the normal approximation.

[21]Alternatively, if you divide the estimated value of any ρ_k by the standard error of ($\sqrt{1/n}$), for sufficiently large n, you will obtain the standard Z value, whose probability can be easily obtained from the standard normal table. Thus for the estimated $\rho_{10} = 0.638$, the Z value is $0.638/0.1066 = 5.98$ (approx.). If the true ρ_{10} were in fact zero, the probability of obtaining a Z value of as much as 5.98 or greater is very small, thus rejecting the hypothesis that the true ρ_{10} is zero.

[22] G. E. P. Box and D. A. Pierce, “Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models,” *Journal of the American Statistical Association*, vol. 65, 1970, pp. 1509–1526.

[23]G. M. Ljung and G. P. E. Box, “On a Measure of Lack of Fit in Time Series Models,” *Biometrika*, vol. 66, 1978, pp. 66–72.

[24]The Q and LB statistics may not be appropriate in every case. For a critique, see Maddala et. al., op. cit., p. 19.

21.9 THE UNIT ROOT TEST

A test of stationarity (or nonstationarity) that has become widely popular over the past several years is the **unit root test.** We will first explain it, then illustrate it and then consider some limitations of this test.

The starting point is the unit root (stochastic) process that we discussed in Section 21.4. We start with

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

where u_t is a white noise error term.

We know that if $\rho = 1$, that is, in the case of the unit root, (21.4.1) becomes a random walk model without drift, which we know is a nonstationary stochastic process. Therefore, why not simply regress Y_t on its (one-period) lagged value Y_{t-1} and find out if the estimated ρ is statistically equal to 1? If it is, then Y_t is nonstationary. This is the general idea behind the unit root test of stationarity.

For theoretical reasons, we manipulate (21.4.1) as follows: Subtract Y_{t-1} from both sides of (21.4.1) to obtain:

$$\begin{aligned} Y_t - Y_{t-1} &= \rho Y_{t-1} - Y_{t-1} + u_t \\ &= (\rho - 1)Y_{t-1} + u_t \end{aligned} \tag{21.9.1}$$

which can be alternatively written as:

$$\Delta Y_t = \delta Y_{t-1} + u_t \tag{21.9.2}$$

where $\delta = (\rho - 1)$ and Δ, as usual, is the first-difference operator.

In practice, therefore, instead of estimating (21.4.1), we estimate (21.9.2) and test the (null) hypothesis that $\delta = 0$. If $\delta = 0$, then $\rho = 1$, that is we have a unit root, meaning the time series under consideration is nonstationary.

Before we proceed to estimate (21.9.2), it may be noted that if $\delta = 0$, (21.9.2) will become

$$\Delta Y_t = (Y_t - Y_{t-1}) = u_t \tag{21.9.3}$$

Since u_t is a white noise error term, it is stationary, which means that the first differences of a random walk time series are stationary, a point we have already made before.

Now let us turn to the estimation of (21.9.2). This is simple enough; all we have to do is to take the first differences of Y_t and regress them on Y_{t-1} and see if the estimated slope coefficient in this regression ($= \hat{\delta}$) is zero or not. If it is zero, we conclude that Y_t is nonstationary. But if it is negative, we conclude that Y_t is stationary.[25] The only question is which test we use to

[25]Since $\delta = (\rho - 1)$, for stationarity ρ must be less than one. For this to happen δ must be negative.

find out if the estimated coefficient of Y_{t-1} in (21.9.2) is zero or not. You might be tempted to say, why not use the usual t test? Unfortunately, under the null hypothesis that $\delta = 0$ (i.e., $\rho = 1$), the t value of the estimated coefficient of Y_{t-1} does not follow the t distribution even in large samples; that is, it does not have an asymptotic normal distribution.

What is the alternative? Dickey and Fuller have shown that under the null hypothesis that $\delta = 0$, the estimated t value of the coefficient of Y_{t-1} in (21.9.2) follows the **τ (tau) statistic.**[26] These authors have computed the critical values of the *tau statistic* on the basis of Monte Carlo simulations. A sample of these critical values is given in **Appendix D**, Table D.7. The table is limited, but MacKinnon has prepared more extensive tables, which are now incorporated in several econometric packages.[27] In the literature the **tau statistic or test** is known as the **Dickey–Fuller (DF) test,** in honor of its discoverers. Interestingly, if the hypothesis that $\delta = 0$ is rejected (i.e., the time series is stationary), we can use the usual (Student's) t test.

The actual procedure of implementing the DF test involves several decisions. In discussing the nature of the unit root process in Sections 21.4 and 21.5, we noted that a random walk process may have no drift, or it may have drift or it may have both deterministic and stochastic trends. To allow for the various possibilities, the DF test is estimated in three different forms, that is, under three different null hypotheses.

Y_t is a random walk: $$\Delta Y_t = \delta Y_{t-1} + u_t \qquad \textbf{(21.9.2)}$$

Y_t is a random walk with drift: $$\Delta Y_t = \beta_1 + \delta Y_{t-1} + u_t \qquad \textbf{(21.9.4)}$$

Y_t is a random walk with drift around a stochastic trend: $$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + u_t \qquad \textbf{(21.9.5)}$$

where t is the time or trend variable. In each case, the *null hypothesis* is that $\delta = 0$; that is, there is a unit root—the time series is nonstationary. The alternative hypothesis is that δ is less than zero; that is, the time series is stationary.[28] If the null hypothesis is rejected, it means that Y_t is a stationary time series with zero mean in the case of (21.9.2), that Y_t is stationary with a nonzero mean $[= \beta_1/(1-\rho)]$ in the case of (21.9.4), and that Y_t is stationary around a deterministic trend in (21.9.5).

[26]D. A. Dickey and W. A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association,* vol. 74, 1979, pp. 427–431. See also W. A. Fuller, *Introduction to Statistical Time Series,* John Wiley & Sons, New York, 1976.

[27]J. G. MacKinnon, "Critical Values of Cointegration Tests," in R. E. Engle and C. W. J. Granger, eds., *Long-Run Economic Relationships: Readings in Cointegration,* Chap. 13, Oxford University Press, New York, 1991.

[28]We rule out the possibility that $\delta > 0$, because in that case $\rho > 1$, in which case the underlying time series will be explosive.

It is extremely important to note that the critical values of the tau test to test the hypothesis that $\delta = 0$, *are different for each of the preceding three specifications of the DF test,* which can be seen clearly from **Appendix D**, Table D.7. Moreover, if, say, specification (21.9.4) is correct, but we estimate (21.9.2), we will be committing a specification error, whose consequences we already know from Chapter 13. The same is true if we estimate (21.9.4) rather than the true (21.9.5). Of course, there is no way of knowing which specification is correct to begin with. Some trial and error is inevitable, data mining notwithstanding.

The actual estimation procedure is as follows: Estimate (21.9.2), or (21.9.3), or (21.9.4) by OLS; divide the estimated coefficient of Y_{t-1} in each case by its standard error to compute the (τ) tau statistic; and refer to the DF tables (or any statistical package). If the computed absolute value of the tau statistic ($|\tau|$) exceeds the DF or MacKinnon critical tau values, we reject the hypothesis that $\delta = 0$, in which case the time series is stationary. On the other hand, if the computed $|\tau|$ does not exceed the critical tau value, we do not reject the null hypothesis, in which case the time series is nonstationary. Make sure that you use the appropriate critical τ values.

Let us return to the U.S. GDP time series. For this series, the results of the three regressions (21.9.2), (21.9.4), and (21.9.5) are as follows: The dependent variable in each case is $\Delta Y_t = \Delta \text{GDP}_t$

$$\widehat{\Delta\text{GDP}}_t = 0.00576\text{GDP}_{t-1}$$
$$t = (5.7980) \qquad R^2 = -0.0152 \qquad d = 1.34 \tag{21.9.6}$$

$$\widehat{\Delta\text{GDP}}_t = 28.2054 - 0.00136\text{GDP}_{t-1}$$
$$t = (1.1576) \quad (-0.2191) \qquad R^2 = 0.00056 \qquad d = 1.35 \tag{21.9.7}$$

$$\widehat{\Delta\text{GDP}}_t = 190.3857 + 1.4776t - 0.0603\text{GDP}_{t-1}$$
$$t = (1.8389) \quad (1.6109) \quad (-1.6252) \tag{21.9.8}$$
$$R^2 = 0.0305 \qquad d = 1.31$$

Our primary interest here is in the t ($= \tau$) value of the GDP_{t-1} coefficient. The critical 1, 5, and 10 percent τ values for model (21.9.6) are −2.5897, −1.9439, and −1.6177, respectively, and are −3.5064, −2.8947, and −2.5842 for model (21.9.7) and −4.0661, −3.4614, and −3.1567 for model (21.3.8). As noted before, these critical values are different for the three models.

Before we examine the results, we have to decide which of the three models may be appropriate. We should rule out model (21.9.6) because the coefficient of GDP_{t-1}, which is equal to δ is positive. But since $\delta = (\rho - 1)$, a positive δ would imply that $\rho > 1$. Although a theoretical possibility, we rule this case out because in this case the GDP time series would be explosive.[29] That leaves us with models (21.9.7) and (21.9.8). In both cases the estimated δ coefficient is negative, implying that the estimated ρ is less than 1. For these two models, the estimated ρ values are 0.9986 and 0.9397, respectively. The only question now is if these values are statistically significantly below 1 for us to declare that the GDP time series is stationary.

For model (21.9.7) the estimated τ value is −0.2191, which in absolute value is below even the 10 percent critical value of −2.5842. Since, in absolute terms, the former is smaller than the latter, our conclusion is that the GDP time series is not stationary.[30]

The story is the same for model (21.9.8). The computed τ value of −1.6252 is less than even the 10 percent critical τ value of −3.1567 in absolute terms.

Therefore, on the basis of graphical analysis, the correlogram, and the Dickey–Fuller test, the conclusion is that for the quarterly periods of 1970 to 1991, the U.S. GDP time series was nonstationary; i.e., it contained a unit root.

The Augmented Dickey–Fuller (ADF) Test

In conducting the DF test as in (21.9.2), (21.9.4), or (21.9.5), it was assumed that the error term u_t was uncorrelated. But in case the u_t are correlated, Dickey and Fuller have developed a test, known as the **augmented Dickey–Fuller (ADF) test.** This test is conducted by "augmenting" the preceding three equations by adding the lagged values of the dependent variable ΔY_t. To be specific, suppose we use (21.9.5). The ADF test here consists of estimating the following regression:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t \qquad \textbf{(21.9.9)}$$

where ε_t is a pure white noise error term and where $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2})$, $\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$, etc. The number of lagged difference terms to include is often determined empirically, the idea being to include enough terms so that the error term in (21.9.9) is serially uncorrelated. In ADF we still test whether $\delta = 0$ and the ADF test follows the same asymptotic distribution as the DF statistic, so the same critical values can be used.

[29] More technically, since (21.9.2) is a first-order difference equation, the so-called stability condition requires that $|\rho| < 1$.

[30] Another way of stating this is that the computed τ value should be more negative than the critical τ value, which is not the case here. Hence the conclusion stays. Since in general δ is expected to be negative, the estimated τ statistic will have a negative sign. Therefore, a large negative τ value is generally an indication of stationarity.

To give a glimpse of this procedure, we estimated (21.9.9) for the GDP series using one lagged difference of GDP; the results were as follows[31]:

$$\widehat{\Delta\text{GDP}}_t = 234.9729 + 1.8921t - 0.0786\text{GDP}_{t-1} + 0.3557\Delta\text{GDP}_{t-1}$$
$$t = \quad (2.3833) \quad (2.1522) \quad (-2.2152) \qquad (3.4647)$$
$$R^2 = 0.1526 \qquad d = 2.0858 \tag{21.9.10}$$

The t ($= \tau$) value of the GDP_{t-1} coefficient ($= \delta$) is -2.2152, but this value in absolute terms is much less than even the 10 percent critical τ value of -3.1570, again suggesting that even after taking care of possible autocorrelation in the error term, the GDP series is nonstationary.

Testing the Significance of More Than One Coefficient: The *F* Test

Suppose we estimate model (21.9.5) and test the hypothesis that $\beta_1 = \beta_2 = 0$, that is, the model is RWM without drift and trend. To test this joint hypothesis, we can use the *restricted F* test discussed in Chapter 8. That is, we estimate (21.9.5) (the unrestricted regression) and estimate (21.9.5), dropping the intercept and trend. Then we use the restricted F test as shown in Eq. (8.7.9), except that we cannot use the conventional F table to get the critical F values. As they did with the τ statistic, Dickey and Fuller have developed critical F values for this situation, a sample of which is given in **Appendix D**, Table D.7. An example is presented in exercise 21.27.

The Phillips–Perron (PP) Unit Root Tests[32]

An important assumption of the DF test is that the error terms u_t are independently and identically distributed. The ADF test adjusts the DF test to take care of possible serial correlation in the error terms by adding the lagged difference terms of the regressand. Phillips and Perron use *nonparametric statistical methods* to take care of the serial correlation in the error terms without adding lagged difference terms. Since the asymptotic distribution of the PP test is the same as the ADF test statistic, we will not pursue this topic here.

A Critique of the Unit Root Tests[33]

We have discussed several unit root tests and there are several more. The question is: Why are there so many unit root tests? The answer lies in the

[31]Higher-order lagged differences were considered but they were insignificant.
[32]P. C. B. Phillips and P. Perron, "Testing for a Unit Root in Time Series Regression," *Biometrika*, vol. 75, 1988, pp. 335–346. The PP test is now included in several software packages.
[33]For detailed discussion, see Terrence C. Mills, op. cit., pp. 87–88.

size and **power** of these tests. By size of a test we mean the level of significance (i.e., the probability of committing a Type I error) and by power of a test we mean the probability of rejecting the null hypothesis when it is false. The power of a test is calculated by subtracting the probability of a Type II error from 1; Type II error is the probability of accepting a false null hypothesis. The maximum power is 1. Most unit root tests are based on the null hypothesis that the time series under consideration has a unit root; that is, it is nonstationary. The alternative hypothesis is that the time series is stationary.

Size of Test. You will recall from Chapter 13 the distinction we made between the nominal and the true levels of significance. The DF test is sensitive to the way it is conducted. Remember that we discussed three varieties of the DF test: (1) a pure random walk, (2) a random walk with drift, and (3) a random walk with drift and trend. If, for example, the true model is (1) but we estimate (2), and conclude that, say, on the 5 percent level that the time series is stationary, this conclusion may be wrong because the true level of significance in this case is much larger than 5 percent.[34] The size distortion could also result from excluding moving average (MA) components from the model (on moving average, see Chapter 22).

Power of Test. Most tests of the DF type have low power; that is, they tend to accept the null of unit root more frequently than is warranted. That is, these tests may find a unit root even when none exists. There are several reasons for this. *First,* the power depends on the (time) *span* of the data more than mere size of the sample. For a given sample size n, the power is greater when the span is large. Thus, unit root test(s) based on 30 observations over a span of 30 years may have more power than that based on, say, 100 observations over a span of 100 days. *Second,* if $\rho \approx 1$ but not exactly 1, the unit root test may declare such a time series nonstationary. *Third,* these types of tests assume a single unit root; that is, they assume that the given time series is $I(1)$. But if a time series is integrated of order higher than 1, say, $I(2)$, there will be more than one unit root. In the latter case one may use the **Dickey–Pantula test.**[35] *Fourth,* if there are structural breaks in a time series (see the chapter on dummy variables) due to, say, the OPEC oil embargoes, the unit root tests may not catch them.

In applying the unit root tests one should therefore keep in mind the limitations of the tests. Of course, there have been modifications of these tests by Perron and Ng, Elliot, Rothenberg and Stock, Fuller, and Leybounre.[36] Because of this, Maddala and Kim advocate that the traditional DF, ADF,

[34]For a Monte Carlo experiment about this, see Charemza et al., op. cit., p. 114.

[35]D. A. Dickey and S. Pantula, "Determining the Order of Differencing in Autoregressive Processes," *Journal of Business and Economic Statistics,* vol. 5, 1987, pp. 455–461.

[36]A discussion of these tests can be found in Maddala et al., op. cit., Chap. 4.

and PP tests should be discarded. As econometric software packages incorporate the new tests, that may very well happen. But it should be added that as yet there is no uniformly powerful test of the unit root hypothesis.

21.10 TRANSFORMING NONSTATIONARY TIME SERIES

Now that we know the problems associated with nonstationary time series, the practical question is what to do. To avoid the spurious regression problem that may arise from regressing a nonstationary time series on one or more nonstationary time series, we have to transform nonstationary time series to make them stationary. The transformation method depends on whether the time series are difference stationary (DSP) or trend stationary (TSP). We consider each of these methods in turn.

Difference-Stationary Processes

If a time series has a unit root, the first differences of such time series are stationary.[37] Therefore, the solution here is to take the first differences of the time series.

Returning to our U.S. GDP time series, we have already seen that it has a unit root. Let us now see what happens if we take the first differences of the GDP series.

Let $\Delta\text{GDP}_t = (\text{GDP}_t - \text{GDP}_{t-1})$. For convenience, let $D_t = \Delta\text{GDP}_t$. Now consider the following regression:

$$\begin{aligned} \widehat{\Delta D_t} &= 16.0049 \quad - \quad 0.06827 D_{t-1} \\ t &= (3.6402) \qquad (-6.6303) \\ & \qquad R^2 = 0.3435 \qquad d = 2.0344 \end{aligned} \tag{21.10.1}$$

The 1 percent critical DF τ value is -3.5073. Since the computed $\tau\ (= t)$ is more negative than the critical value, we conclude that the first-differenced GDP is stationary; that is, it is $I(0)$. It is as shown in Figure 21.9. If you compare Figure 21.9 with Figure 21.1, you will see the obvious difference between the two.

Trend-Stationary Process

As we have seen in Figure 21.5, a TSP is stationary around the trend line. Hence, the simplest way to make such a time series stationary is to regress it on time and the residuals from this regression will then be stationary.

[37] If a time series is $I(2)$, it will contain two unit roots, in which case we will have to difference it twice. If it is $I(d)$, it has to be differenced d times, where d is any integer.

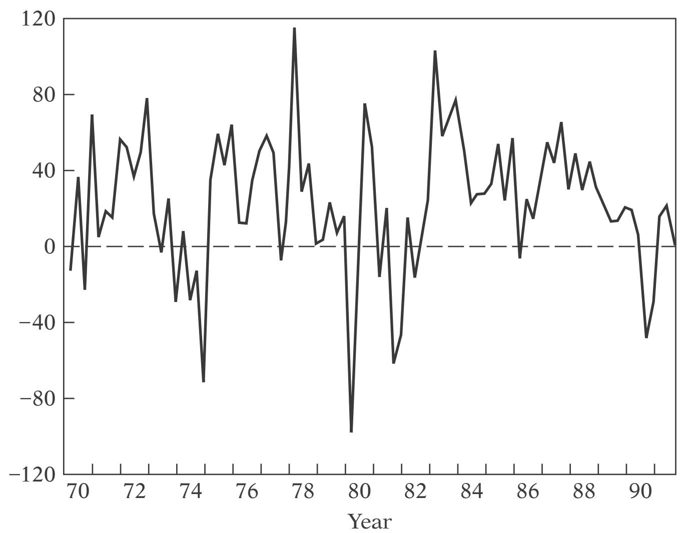


FIGURE 21.9 First differences of U.S. GDP, 1970–1991 (quarterly).

In other words, run the following regression:

$$Y_t = \beta + \beta_2 t + u_t \qquad \textbf{(21.10.2)}$$

where Y_t is the time series under study and where t is the trend variable measured chronologically.

Now

$$\hat{u}_t = (Y_t - \hat{\beta}_1 - \hat{\beta}_2 t) \qquad \textbf{(21.10.3)}$$

will be stationary. $\hat{u}_t$ is known as a (linearly) **detrended time series.**

It is important to note that the trend may be nonlinear. For example, it could be

$$Y_t = \beta_1 + \beta_2 t + \beta_3 t^2 + u_t \qquad \textbf{(21.10.4)}$$

which is a quadratic trend series. If that is the case, the residuals from (21.10.4) will now be (quadratically) detrended time series.

It should be pointed out that if a time series is DSP but we treat it as TSP, this is called **underdifferencing.** On the other hand, if a time series is TSP but we treat it as DSP, this is called **overdifferencing.** The consequences of these types of specification errors can be serious, depending on how one handles the serial correlation properties of the resulting error terms.[38]

In passing it may be noted that most macroeconomic time series are DSP rather than TSP.

[38]For a detailed discussion of this, see Maddala et al., op. cit., Sec. 2.7.

21.11 COINTEGRATION: REGRESSION OF A UNIT ROOT TIME SERIES ON ANOTHER UNIT ROOT TIME SERIES

We have warned that the regression of a nonstationary time series on another nonstationary time series may produce a spurious regression. Let us suppose that we consider the PCE and PDI time series given in Table 21.1. Subjecting these time series individually to unit root analysis, you will find that they both are $I(1)$; that is, they contain a unit root. Suppose, then, that we regress PCE on PDI as follows:

$$\text{PCE}_t = \beta_1 + \beta_2\text{PDI}_t + u_t \quad \textbf{(21.11.1)}$$

Let us write this as:

$$u_t = \text{PCE}_t - \beta_1 - \beta_2\text{PDI}_t \quad \textbf{(21.11.2)}$$

Suppose we now subject u_t to unit root analysis and find that it is stationary; that is, it is $I(0)$. This is an interesting situation, for although PCE_t and PDI_t are individually $I(1)$, that is, they have stochastic trends, their linear combination (21.11.2) is $I(0)$. So to speak, the linear combination cancels out the stochastic trends in the two series. If you take consumption and income as two $I(1)$ variables, savings defined as (income − consumption) could be $I(0)$. As a result, a regression of consumption on income as in (21.11.1) would be meaningful (i.e., not spurious). In this case we say that the two variables are **cointegrated.** Economically speaking, two variables will be cointegrated if they have a long-term, or equilibrium, relationship between them. Economic theory is often expressed in equilibrium terms, such as Fisher's quantity theory of money or the theory of purchasing parity (PPP), just to name a few.

In short, provided we check that the residuals from regressions like (21.11.1) are $I(0)$ or stationary, the traditional regression methodology (including the t and F tests) that we have considered extensively is applicable to data involving (nonstationary) time series. The valuable contribution of the concepts of unit root, cointegration, etc. is to force us to find out if the regression residuals are stationary. As Granger notes, "A test for cointegration can be thought of as a pre-test to avoid 'spurious regression' situations."[39]

In the language of cointegration theory, a regression such as (21.11.1) is known as a **cointegrating regression** and the slope parameter β_2 is known as the **cointegrating parameter.** The concept of cointegration can be extended to a regression model containing k regressors. In this case we will have k cointegrating parameters.

Testing for Cointegration

A number of methods for testing cointegration have been proposed in the literature. We consider here two comparatively simple methods: (1) the DF

[39]C. W. J. Granger, "Developments in the Study of Co-Integrated Economic Variables," *Oxford Bulletin of Economics and Statistics,* vol. 48, 1986, p. 226.

or ADF unit root test on the residuals estimated from the cointegrating regression and (2) the cointegrating regression Durbin–Watson (CRDW) test.[40]

Engle–Granger (EG) or Augmented Engle–Granger (AEG) Test. We already know how to apply the DF or ADF unit root tests. All we have to do is estimate a regression like (21.11.1), obtain the residuals, and use the DF or ADF tests.[41] There is one precaution to exercise, however. Since the estimated u_t are based on the *estimated* cointegrating parameter β_2, the DF and ADF critical significance values are not quite appropriate. Engle and Granger have calculated these values, which can be found in the references.[42] Therefore, the DF and ADF tests in the present context are known as **Engle–Granger (EG)** and **augmented Engle–Granger (AEG)** tests. However, several software packages now present these critical values along with other outputs.

Let us illustrate these tests. We first regressed PCE on PDI and obtained the following regression:

$$\begin{aligned}\widehat{\text{PCE}}_t &= -171.4412 + 0.9672\text{PDI}_t \\ t &= \quad (-7.4808) \quad (119.8712) \\ &R^2 = 0.9940 \qquad d = 0.5316\end{aligned} \tag{21.11.3}$$

Since PCE and PDI are individually nonstationary, there is the possibility that this regression is spurious. But when we performed a unit root test on the residuals obtained from (21.11.3), we obtained the following results:

$$\begin{aligned}\widehat{\Delta\hat{u}_t} &= -0.2753\hat{u}_{t-1} \\ t &= (-3.7791) \\ &R^2 = 0.1422 \qquad d = 2.2775\end{aligned} \tag{21.11.4}$$

The Engle–Granger 1 percent critical τ value is -2.5899. Since the computed $\tau\,(=t)$ value is much more negative than this, our conclusion is that the residuals from the regression of PCE on PDI are $I(0)$; that is, they are

[40]There is this difference between tests for unit roots and tests for cointegration. As David A. Dickey, Dennis W. Jansen, and Daniel I. Thornton observe, “Tests for unit roots are performed on univariate [i.e., single] time series. In contrast, cointegration deals with the relationship among a group of variables, where (unconditionally) each has a unit root.” See their article, “A Primer on Cointegration with an Application to Money and Income,” *Economic Review,* Federal Reserve Bank of St. Louis, March-April 1991, p. 59. As the name suggests, this article is an excellent introduction to cointegration testing.

[41]If PCE and PDI are not cointegrated, any linear combination of them will be nonstationary and, therefore, the u_t will also be nonstationary.

[42]R. F. Engle and C. W. Granger, “Co-integration and Error Correction: Representation, Estimation and Testing,” *Econometrica*, vol. 55, 1987, pp. 251–276.

stationary. Hence, (21.11.3) is a cointegrating regression and this regression is not spurious, even though individually the two variables are nonstationary. One can call (21.11.3) the **static** or **long run** consumption function and interpret its parameters as long run parameters. Thus, 0.9672 represents the long-run, or equilibrium, marginal propensity to consumer (MPC).

Cointegrating Regression Durbin–Watson (CRDW) Test. An alternative, and quicker, method of finding out whether PCE and PDI are cointegrated is the CRDW test, whose critical values were first provided by Sargan and Bhargava.[43] In CRDW we use the Durbin–Watson d obtained from the cointegrating regression, such as $d = 0.5316$ given in (21.11.3). But now the null hypothesis is that $d = 0$ rather than the standard $d = 2$. This is because in Chapter 12 we observed that $d \approx 2(1 - \hat{\rho})$, so if there is to be a unit root, the estimated ρ will be about 1, which implies that d will be about zero.

On the basis of 10,000 simulations formed from 100 observations each, the 1, 5, and 10 percent critical values to test the hypothesis that the true $d = 0$ are 0.511, 0.386, and 0.322, respectively. Thus, if the computed d value is smaller than, say, 0.511, we reject the null hypothesis of cointegration at the 1 percent level. In our example, the value of 0.5316 is above this critical value, suggesting that PCE and PDI are cointegrated, thus reinforcing the finding on the basis of the EG test.[44]

To sum up, our conclusion, based on both the EG and CRDW tests, is that PCE and PDI are cointegrated.[45] Although they individually exhibit random walks, there seems to be a stable long-run relationship between them; they will not wander away from each other, which is evident from Figure 21.1.

Cointegration and Error Correction Mechanism (ECM)

We just showed that PCE and PDI are cointegrated; that is, there is a long-term, or equilibrium, relationship between the two. Of course, in the short run there may be disequilibrium. Therefore, one can treat the error term in (21.11.2) as the "equilibrium error." And we can use this error term to tie the short-run behavior of PCE to its long-run value. The **error correction mechanism (ECM)** first used by Sargan[46] and later popularized by Engle

[43] J. D. Sargan and A. S. Bhargava, "Testing Residuals from Least-Squares Regression for being Generated by the Gaussian Random Walk," *Econometrica*, vol. 51, 1983, pp. 153–174.

[44] There is considerable debate about the superiority of CRDW over DF, which can be found in the references. The debate revolves around the power of the two statistics, that is, the probability of not committing a Type II error. Engle and Granger, for example, prefer the ADF to the CRDW test.

[45] The EG and CRDW tests are now supplemented (supplanted?) by more powerful tests developed by Johansen. But the discussion of the **Johansen method** is beyond the scope of this book because the mathematics involved is quite complex, although several software packages now use the Johansen method.

[46] J. D. Sargan, "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in K. F. Wallis and D. F. Hendry, eds., *Quantitative Economics and Econometric Analysis*, Basil Blackwell, Oxford, U.K., 1984.

and Granger corrects for disequilibrium. An important theorem, known as the **Granger representation theorem,** states that if two variables Y and X are cointegrated, then the relationship between the two can be expressed as ECM. To see what this means, let us revert to our PCE–PDI example. Now consider the following model:

$$\Delta \text{PCE}_t = \alpha_0 + \alpha_1 \Delta \text{PDI}_t + \alpha_2 u_{t-1} + \varepsilon_t \qquad \textbf{(21.11.5)}$$

where Δ as usual denotes the first difference operator, ε_t is a random error term, and $u_{t-1} = (\text{PCE}_{t-1} - \beta_1 - \beta_2 \text{PDI}_{t-1})$, that is, the one-period lagged value of the error from the cointegrating regression (21.11.1).

ECM equation (21.11.5) states that ΔPCE depends on ΔPDI and also on the equilibrium error term.[47] If the latter is nonzero, then the model is out of equilibrium. Suppose ΔPDI is zero and u_{t-1} is positive. This means PCE_{t-1} is too high to be in equilibrium, that is, PCE_{t-1} is above its equilibrium value of $(\alpha_0 + \alpha_1 \text{PDI}_{t-1})$. Since α_2 is expected to be negative, the term $\alpha_2 u_{t-1}$ is negative and, therefore, ΔPCE_t will be negative to restore the equilibrium. That is, if PCE_t is above its equilibrium value, it will start falling in the next period to correct the equilibrium error; hence the name ECM. By the same token, if u_{t-1} is negative (i.e., PCE is below its equilibrium value), $\alpha_2 u_{t-1}$ will be positive, which will cause ΔCPE_t to be positive, leading PCE_t to rise in period t. Thus, the absolute value of α_2 decides how quickly the equilibrium is restored. In practice, we estimate u_{t-1} by $\hat{u}_{t-1} = (\text{PCE}_t - \hat{\beta}_1 - \hat{\beta}_2 \text{PDI}_t)$.

Returning to our illustrative example, the empirical counterpart of (21.11.5) is:

$$\begin{aligned} \widehat{\Delta \text{PCE}}_t &= 11.6918 \ + \ 0.2906 \Delta \text{PDI}_t - \ 0.0867 \hat{u}_{t-1} \\ t &= (5.3249) \quad (4.1717) \qquad (-1.6003) \qquad \textbf{(21.11.6)} \\ & \qquad R^2 = 0.1717 \qquad d = 1.9233 \end{aligned}$$

Statistically, the equilibrium error term is zero, suggesting that PCE adjusts to changes in PDI in the same time period. As (21.11.6) shows, short-run changes in PDI have a positive impact on short-run changes in personal consumption. One can interpret 0.2906 as the short-run marginal propensity to consume (MPC); the long-run MPC is given by the estimated (static) equilibrium relation (21.11.3) as 0.9672.

Before we conclude this section, the caution sounded by S. G. Hall is worth remembering:

> While the concept of cointegration is clearly an important theoretical underpinning of the error correction model there are still a number of problems surrounding its

[47] The following discussion is based on Gary Koop, op. cit., pp. 159–160 and Kerry Peterson, op. cit., Sec. 8.5.

practical application; the critical values and small sample performance of many of these tests are unknown for a wide range of models; informed inspection of the correlogram may still be an important tool.[48]

21.12 SOME ECONOMIC APPLICATIONS

We conclude this chapter by considering some concrete examples.

EXAMPLE 21.1

M1 MONTHLY MONEY SUPPLY IN THE UNITED STATES, JANUARY 1951 TO SEPTEMBER 30, 1999

Figure 21.10 shows the M1 money supply for the United States from January 1951 to September 30, 1999. From our knowledge of stationarity, it seems that the M1 money supply time series is nonstationary, which can be confirmed by unit root analysis. (*Note:* to save space, we have not given the actual data, which can be obtained from the Federal Reserve Board

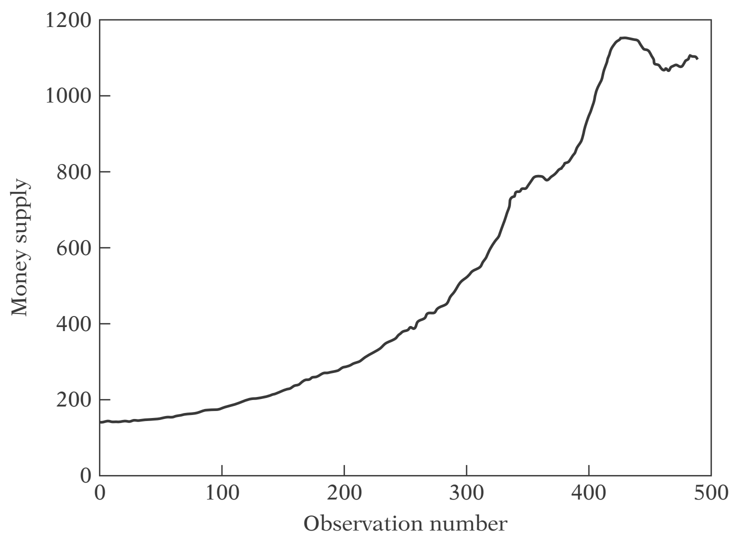


FIGURE 21.10 U.S. money supply over 1951:01 to 1999:09.

(*Continued*)

[48]S. G. Hall, "An Application of the Granger and Engle Two-Step Estimation Procedure to the United Kingdom Aggregate Wage Data," *Oxford Bulletin of Economics and Statistics,* vol. 48, no. 3, August 1986, p. 238. See also John Y. Campbell and Pierre Perron, "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots," NBER (National Bureau of Economic Research) *Macroeconomics Annual 1991,* pp. 141–219.

EXAMPLE 21.1 (*Continued*)

or the Federal Reserve Bank of St. Louis.)

$$\Delta\hat{M}_t = 0.2618 + 0.0159t - 0.0044M_{t-1}$$
$$t = (0.7919) \quad (4.4227) \quad (-3.0046) \qquad \textbf{(21.12.1)}$$
$$R^2 = 0.0670 \qquad d = 0.7172$$

The 1, 5, and 10 percent critical τ values are −3.9811, −3.4210, and −3.1329. Since the t value of −3.0046 is less negative than any of these critical values, the conclusion is that the M1 time series is nonstationary; that is, it contains a unit root or it is $I(1)$. Even when several lagged values of ΔM_t (à la ADF) were introduced, the conclusion did not change. On the other hand, the first differences of the M1 money supply were found to be stationary (check this out).

EXAMPLE 21.2

THE U.S./U.K. EXCHANGE RATE: JANUARY 1, 1973, TO OCTOBER 10, 1996

Figure 21.11 gives the graph of the ($/£) exchange rate from January 1973 to October 1996, for a total of 286 observations. By now you should be able to spot this time series as nonstationary. Carrying out the unit root tests, we obtained the following τ statistics: −1.2749 (no intercept, no trend), −1.7710 (intercept), and −1.6269 (intercept and trend). Each of these statistics, in absolute value, was less than its critical τ value from the appropriate DF tables, thus confirming the graphical impression that the U.S./U.K. exchange rate time series is nonstationary.

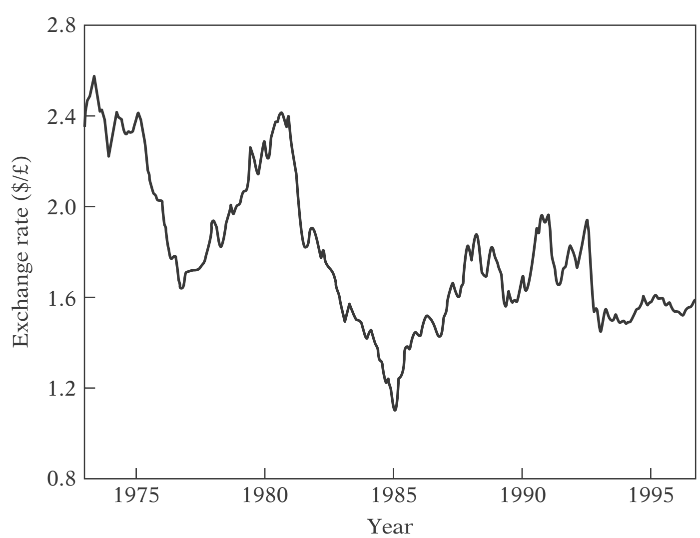


FIGURE 21.11 U.S./U.K. exchange rate: January 1973 to October 1996.

EXAMPLE 21.3

U.S. CONSUMER PRICE INDEX (CPI), JANUARY 1947 TO JANUARY 2000

Figure 21.12 shows the U.S. CPI from January 1947 to January 2000 for a total of 649 observations. The CPI series, like the M1 series considered previously, shows a sustained upward trend. The unit root exercise gave the following results:

$$\widehat{\Delta \text{CPI}}_t = -0.0094 + 0.00051t - 0.00066\text{CPI}_{t-1} + 0.5473\Delta\text{CPI}_{t-1}$$
$$t = (-0.6538) \quad (4.3431) \quad (-1.5472) \quad (16.4448) \qquad \textbf{(21.12.2)}$$
$$R^2 = 0.5177 \qquad d = 2.1410$$

The t ($= \tau$) value of CPI_{t-1} is -1.5472. The 10 percent critical value is -3.1317. Since, in absolute terms, the computed τ is less than the critical τ, the conclusion is that CPI is not a stationary time series. We can characterize it as having a stochastic trend (why?). However, if you take the first differences of the CPI series, you will find them to be stationary. Hence CPI is a difference-stationary (DS) time series.

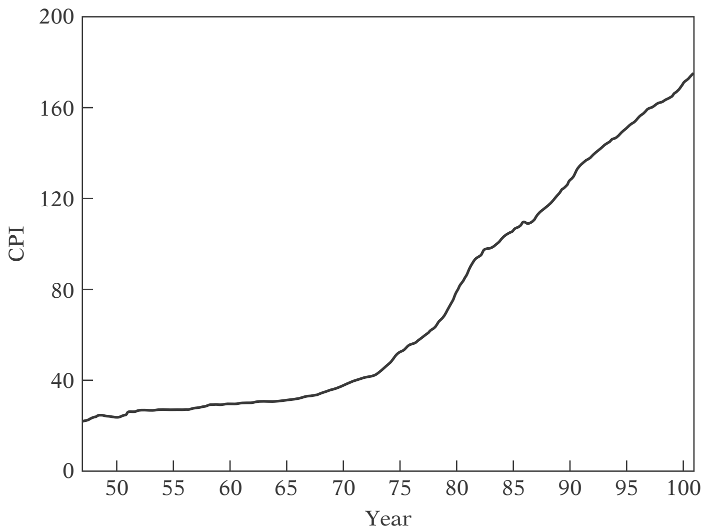


FIGURE 21.12 U.S. CPI, January 1947 to January 2000.

EXAMPLE 21.4

ARE 3-MONTH AND 6-MONTH TREASURY BILL RATES COINTEGRATED?

Figure 21.13 plots (constant maturity) 3-month and 6-month U.S. Treasury bill (T bill) rates from January 1982 to June 2001, for a total of 234 observations. Does the graph show that the two rates are cointegrated; that is, is there an equilibrium relationship between the two? From financial theory, we would expect that to be the case, otherwise arbitrageurs will exploit any discrepancy between the short and the long rates. First of all, let us see if the two time series are stationary.

(Continued)

EXAMPLE 20.1 (*Continued*)

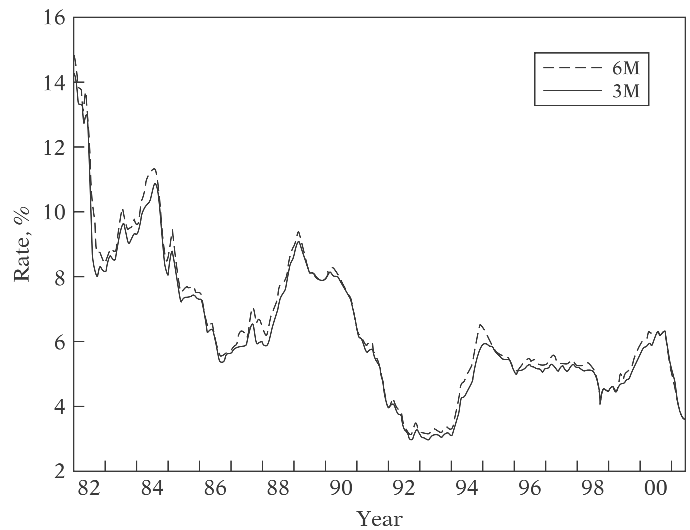


FIGURE 21.13
Three- and six-month Treasury bill rates (constant maturity).

On the basis of the pure random walk model (i.e., no intercept, no trend), both the rates were stationary. Including intercept, trend, and one lagged difference, the results suggested that the two rates might be trend stationary; the trend coefficient in both cases was negative and significant at about the 7 percent level. So, depending on which results we accept, the two rates are either stationary or trend stationary.

Regressing the 6-month T bill rate (TB6) on the 3 month T-bill rate, we obtained the following regression.

$$\widehat{\text{TB6}}_t = -0.0456 + 1.0466\text{TB3}_t$$
$$t = (-1.1207) \quad (171.6239) \qquad R^2 = 0.9921 \qquad d = 0.4055 \qquad \textbf{(21.12.3)}$$

Applying the unit root test to the residuals from the preceding regression, we found that the residuals were stationary, suggesting that the 3- and 6-month T bill rates were cointegrated. Using this knowledge, we obtained the following error correction model (ECM):

$$\Delta\widehat{\text{TB6}}_t = -0.0067 + 0.9360\Delta\text{TB3}_t - 0.2030\hat{u}_{t-1}$$
$$t = (-0.8662) \quad (41.9592) \qquad (-5.3837) \qquad \textbf{(21.12.4)}$$
$$R^2 = 0.8852 \qquad d = 1.5604$$

where $\hat{u}_{t-1}$ is the lagged value of the error correction term from the preceding period. As these results show, 0.20 of the discrepancy in the two rates in the previous month is eliminated this month.[49] Besides, short-run changes in the 3-month T bill rate are quickly reflected in the 6-month T bill rate, as the slope coefficient between the two is 0.9360. This should not be a surprising finding in view of the efficiency of the U.S. money markets.

[49]Since both T bill rates are in percent form, this would suggest that if the 6-month TB rate was higher than the 3-month TB rate more than expected a priori in the last month, this month it will be reduced by 0.20 percentage points to restore the long-run relationship between the two interest rates. For the underlying theory about the relationship between short- and long-run interest rates, see any money and banking textbook and read up on the term structure of interest rates.