Cases 2 or 3 apply. If the variables are cointegrated, the results of Chapter 6 apply. The remainder of this chapter considers the formal test procedures for the presence of unit roots and/or deterministic time trends.

2. DICKEY–FULLER TESTS

The last section outlined a simple procedure to determine whether $a_1 = 1$ in the model $y_t = a_1 y_{t-1} + \epsilon_t$. Begin by subtracting y_{t-1} from each side of the equation in order to write the equivalent form: $\Delta y_t = \gamma y_{t-1} + \epsilon_t$, where $\gamma = a_1 - 1$. Of course, testing the hypothesis $a_1 = 1$ is equivalent to testing the hypothesis $\gamma = 0$. Dickey and Fuller (1979) actually consider three different regression equations that can be used to test for the presence of a unit root:

$$\Delta y_t = \gamma y_{t-1} + \epsilon_t \tag{4.9}$$
$$\Delta y_t = a_0 + \gamma y_{t-1} + \epsilon_t \tag{4.10}$$
$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \epsilon_t \tag{4.11}$$

The difference between the three regressions concerns the presence of the deterministic elements a_0 and $a_2 t$. The first is a pure random walk model, the second adds an intercept or *drift* term, and the third includes both a drift and linear time trend.

The parameter of interest in all the regression equations is γ; if $\gamma = 0$, the $\{y_t\}$ sequence contains a unit root. The test involves estimating one (or more) of the equations above using OLS in order to obtain the estimated value of γ and associated standard error. Comparing the resulting t-statistic with the appropriate value reported in the Dickey–Fuller tables allows the researcher to determine whether to accept or reject the null hypothesis $\gamma = 0$.

Recall that in (4.3), the estimate of $y_t = a_1 y_{t-1} + \epsilon_t$ was such that $a_1 = 0.9546$ with a standard error of 0.030. Clearly, the OLS regression in the form $\Delta y_t = \gamma y_{t-1} + \epsilon_t$ will yield an estimate of γ equal to -0.0454 with the same standard error of 0.030. Hence, the associated t-statistic for the hypothesis $\gamma = 0$ is -1.5133 ($-0.0454/0.03 = -1.5133$).

The methodology is precisely the same, regardless of which of the three forms of the equations is estimated. However, be aware that the critical values of the t-statistics do depend on whether an intercept and/or time trend is included in the regression equation. In their Monte Carlo study, Dickey and Fuller (1979) found that the critical values for $\gamma = 0$ depend on the form of the regression and sample size. The statistics labeled τ, τ_μ, and τ_τ are the appropriate statistics to use for Equations (4.9), (4.10), and (4.11), respectively.

Now, look at Table A at the end of this book. With 100 observations, there are three different critical values for the t-statistic $\gamma = 0$. For a regression without the intercept and trend terms ($a_0 = a_2 = 0$), use the section labeled τ. With 100 observations, the critical values for the t-statistic are -1.61, -1.95 and -2.60 at the 10, 5, and 1% significance levels, respectively. Thus, in the hypothetical example with $\gamma = -0.0454$ and a standard error of 0.03 (so that $t = -1.5133$), it is not possible to reject the null of a unit root at conventional significance levels. Note that the appropriate critical values depend on sample size. As in most hypothesis tests, for any given level of significance, the critical values of the t-statistic decrease as sample size increases.

Including an intercept term but not a trend term (only $a_2 = 0$) necessitates the use of the critical values in the section labeled τ_μ. Estimating (4.4) in the form $\Delta y_t = a_0 + \gamma y_{t-1} + \epsilon_t$ necessarily yields a value of γ equal to $(0.9247 - 1) = -0.0753$ with a standard error of 0.037. The appropriate calculation for the τ_μ statistic yields $-0.0753/0.037 = -2.035$. If we read from the appropriate row of Table A, with the same 100 observations, the critical values are −2.58, −2.89, and −3.51 at the 10, 5, and 1% significance levels, respectively. Again, the null of a unit root cannot be rejected at conventional significance levels. Finally, with both intercept and trend, use the critical values in the section labeled τ_τ; now the critical values are −3.45 and −4.04 at the 5 and 1% significance levels, respectively. The equation was not estimated using a time trend; inspection of Figure 4.1 indicates there is little reason to include a deterministic trend in the estimating equation.

As discussed in the next section, these critical values are unchanged if (4.9), (4.10), and (4.11) are replaced by the autoregressive processes:[6]

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t \tag{4.12}$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t \tag{4.13}$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t \tag{4.14}$$

The same τ, τ_μ, and τ_τ statistics are all used to test the hypotheses $\gamma = 0$. Dickey and Fuller (1981) provide three additional F-statistics (called ϕ_1, ϕ_2 and ϕ_3) to test **joint** hypotheses on the coefficients. With (4.10) or (4.13), the null hypothesis $\gamma = a_0 = 0$ is tested using the ϕ_1 statistic. Including a time trend in the regression—so that (4.11) or (4.14) is estimated—the joint hypothesis $a_0 = \gamma = a_2 = 0$ is tested using the ϕ_2 statistic and the joint hypothesis $\gamma = a_2 = 0$ is tested using the ϕ_3 statistic.

The ϕ_1, ϕ_2, and ϕ_3 statistics are constructed in exactly the same way as ordinary F-tests are:

$$\phi_i = \frac{[\text{RSS(restricted)} - \text{RSS(unrestricted)}]\ r}{\text{RSS(unrestricted)}\ (T-k)}$$

where RSS(restricted) and RSS(unrestricted) = the sums of the squared residuals from the restricted and unrestricted models

r = number of restrictions
T = number of usable observations
k = number of parameters estimated in the unrestricted model

Hence, $T - k$ = degrees of freedom in the unrestricted model.

Comparing the calculated value of ϕ_i to the appropriate value reported in Dickey and Fuller (1981) allows you to determine the significance level at which the restriction is binding. The null hypothesis is that the data are generated by the restricted model and the alternative hypothesis is that the data are generated by the unrestricted model. If the restriction is not binding, RSS(restricted) should be close to RSS(unrestricted) and ϕ_i should be small; hence, large values of ϕ_i suggest a binding restriction and rejection of the null hypothesis. Thus, if the calculated value of ϕ_i is smaller than that reported by Dickey and Fuller, you can accept the restricted model (i.e., you do not reject the null hypothesis that the restriction is not binding). If the calculated value of ϕ_i is larger than reported by Dickey and Fuller, you can reject the null hypothesis and conclude that the restriction is binding. The critical values of the three ϕ_i statistics are reported in Table B at the end of this text.

Finally, it is possible to test hypotheses concerning the significance of the drift term a_0 and time trend a_2. Under the null hypothesis $\gamma = 0$, the test for the presence of the time trend in (4.14) is given by the $\tau_{\beta\tau}$ statistic. Thus, this statistic is the test $a_2 = 0$ given that $\gamma = 0$. To test the hypothesis $a_0 = 0$, use the $\tau_{\alpha\tau}$ statistic if you estimate (4.14) and the $\tau_{\alpha\mu}$ statistic if you estimate (4.13). The complete set of test statistics and their critical values for a sample size of 100 are summarized in Table 4.1.

Table 4.1 **Summary of the Dickey–Fuller Tests**

Model	Hypothesis	Test Statistic	Critical values for 95% and 99% Confidence Intervals
$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \epsilon_t$	$\gamma = 0$	τ_τ	−3.45 and − 4.04
	$a_0 = 0$ given $\gamma = 0$	$\tau_{\alpha\tau}$	3.11 and 3.78
	$a_2 = 0$ given $\gamma = 0$	$\tau_{\beta\tau}$	2.79 and 3.53
	$\gamma = a_2 = 0$	ϕ_3	6.49 and 8.73
	$a_0 = \gamma = a_2 = 0$	ϕ_2	4.88 and 6.50
$\Delta y_t = a_0 + \gamma y_{t-1} + \epsilon_t$	$\gamma = 0$	τ_μ	−2.89 and −3.51
	$a_0 = 0$ given $\gamma = 0$	$\tau_{\alpha\mu}$	2.54 and 3.22
	$a_0 = \gamma = 0$	ϕ_1	4.71 and 6.70
$\Delta y_t = \gamma y_{t-1} + \epsilon_t$	$\gamma = 0$	τ	−1.95 and −2.60

Notes: Critical values are for a sample size of 100.

An Example

To illustrate the use of the various test statistics, Dickey and Fuller (1981) use quarterly values of the logarithm of the Federal Reserve Board's Production Index over the 1950:I to 1977:IV period to estimate the following three equations:

$$\Delta y_t = 0.52 + 0.00120t - 0.119y_{t-1} + 0.498\Delta y_{t-1} + \epsilon_t, \quad RSS = 0.056448$$
$$(0.15)\ (0.00034)\ (0.033)\ (0.081) \qquad (4.15)$$
$$\Delta y_t = 0.0054 + 0.447\Delta y_{t-1} + \epsilon_t, \quad RSS = 0.063211$$
$$(0.0025)\ (0.083) \qquad (4.16)$$
$$\Delta y_t = 0.511\Delta y_{t-1} + \epsilon_t, \quad RSS = 0.065966$$
$$(0.079) \qquad (4.17)$$

where RSS = residual sum of squares, and standard errors are in parentheses.

To test the null hypothesis that the data are generated by (4.17) against the alternative that (4.15) is the "true" model, use the ϕ_2 statistic. Dickey and Fuller test the null hypothesis $a_0 = a_2 = \gamma = 0$ as follows. Note that the residual sums of squares of the restricted and unrestricted models are 0.065966 and 0.056448 and the null hypothesis entails three restrictions. With 110 usable observations and four estimated parameters, the unrestricted model contains 106 degrees of freedom. Since 0.056448/106 = 0.000533, the ϕ_2 statistic is given by

$$\phi_2 = (0.065966 - 0.056448)/\,3(0.000533) = 5.95$$

With 110 observations, the critical value of ϕ_2 calculated by Dickey and Fuller is 5.59 at the 2.5% significance level. Hence, it is possible to reject the null hypothesis of a random walk against the alternative that the data contain an intercept and/or a unit root and/or a deterministic time trend (i.e., rejecting $a_0 = a_2 = \gamma = 0$ means that one or more of these parameters does not equal zero).

Dickey and Fuller also test the null hypothesis $a_2 = \gamma = 0$ given the alternative of (4.15). Now if we view (4.16) as the restricted model and (4.15) as the unrestricted model, the ϕ_3 statistic is calculated as

$$\phi_3 = (0.063211 - 0.056448)/\,2(0.000533) = 6.34$$

With 110 observations, the critical value of ϕ_3 is 6.49 at the 5% significance level and 5.47 at the 10% significance level.[7] At the 10% level, they reject the null hypothesis. However, at the 5% level, the calculated value of ϕ_3 is smaller than the critical value; they do not reject the null hypothesis that the data contain a unit root and/or deterministic time trend.

To compare with the τ_τ test (i.e., the hypothesis that only $\gamma = 0$) note that

$$\tau_\tau = -0.119/0.033 = -3.61.$$

which rejects the null of a unit root at the 5% level.

3. EXTENSIONS OF THE DICKEY–FULLER TEST

Not all time-series processes can be well represented by the first-order autoregressive process $\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \epsilon_t$. It is possible to use the Dickey–Fuller tests in higher-order equations such as (4.12), (4.13), and (4.14). Consider the pth-order autoregressive process:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \cdots + a_{p-2} y_{t-p+2} + a_{p-1} y_{t-p+1} + a_p y_{t-p} + \epsilon_t \qquad (4.18)$$

To best understand the methodology of the **augmented Dickey–Fuller** test, add and subtract $a_p y_{t-p+1}$ to obtain:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \cdots + a_{p-2} y_{t-p+2} + (a_{p-1} + a_p) y_{t-p+1} - a_p \Delta y_{t-p+1} + \epsilon_t$$

Next, add and subtract $(a_{p-1} + a_p) y_{t-p+2}$ to obtain

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \cdots - (a_{p-1} + a_p)\Delta y_{t-p+2} - a_p \Delta y_{t-p+1} + \epsilon_t$$

Continuing in this fashion, we get

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t$$

where $\gamma = -\left(1 - \sum_{i=1}^{p} a_i\right)$

$$\beta_i = \sum_{j=i}^{p} a_j \qquad (4.19)$$

In (4.19), the coefficient of interest is γ; if $\gamma = 0$, the equation is entirely in first differences and so has a unit root. We can test for the presence of a unit root using the same Dickey–Fuller statistics discussed above. Again, the appropriate statistic to use depends on the deterministic components included in the regression equation. Without an intercept or trend, use the τ statistic; with only the intercept, use the τ_μ statistic; and with both an intercept and trend, use the τ_τ statistic. It is worthwhile pointing out that the results here are perfectly consistent with our study of difference equations in Chapter 1. If the coefficients of a difference equation sum to 1, *at least* one characteristic root is unity. Here, if $\Sigma a_i = 1$, $\gamma = 0$ and the system has a unit root.

Note that the Dickey–Fuller tests assume that the errors are independent and have a constant variance. This raises four important problems related to the fact that we do not know the true data-generating process. First, the true data-generating process may contain both autoregressive and moving average components. We need to know how to conduct the test if the order of the moving average terms (if

Determination of the Deterministic Regressors

Unless the researcher knows the actual data-generating process, there is a question concerning whether it is most appropriate to estimate (4.12), (4.13) or (4.14). It might seem reasonable to test the hypothesis $\gamma = 0$ using the most general of the models, that is,

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t \qquad (4.35)$$

After all, if the true process is a random walk process, this regression should find that $a_0 = \gamma = a_2 = 0$. One problem with this line of reasoning is that the presence of the additional estimated parameters reduces degrees of freedom and the power of

the test. Reduced power means that the researcher will conclude that the process contains a unit root when, in fact, none is present. The second problem is that the appropriate statistic (i.e., the τ, τ_μ, and τ_τ) for testing $\gamma = 0$ depends on which regressors are included in the model. As you can see by examining the three Dickey–Fuller tables, for a given significance level, the confidence intervals around $\gamma = 0$ dramatically expand if a drift and time trend are included in the model. This is quite different from the case in which $\{y_t\}$ is stationary. The distribution of the t-statistic does not depend on the presence of the other regressors when stationary variables are used.

The point is that it is important to use a regression equation that mimics the actual data-generating process. If we inappropriately omit the intercept or time trend, the power of the test can go to zero.[10] For example, if as in (4.35), the data-generating process includes a trend, omitting the term $a_2 t$ imparts an upward bias in the estimated value of γ. On the other hand, extra regressors increase the absolute value of the critical values so that you may fail to reject the null of a unit root.

To illustrate the problem, suppose that the time series $\{y_t\}$ is assumed to be generated by the random walk plus drift process:

$$y_t = a_0 + a_1 y_{t-1} + \epsilon_t, \qquad a_0 \neq 0 \text{ and } a_1 = 1 \qquad (4.36)$$

where the initial condition y_0 is given and $t = 1, 2, \ldots, T$.

If there is no drift, it is inappropriate to include the intercept term since the power of the Dickey–Fuller test is reduced. When the drift is actually in the data-generating process, omitting a_0 from the estimating equation also reduces the power of the test in finite samples. How do you know whether to include a drift or time trend in performing the tests? The key problem is that *the tests for unit roots are conditional on the presence of the deterministic regressors and tests for the presence of the deterministic regressors are conditional on the presence of a unit root.* Campbell and Perron (1991) report the following results concerning unit root tests:

1. When the estimated regression includes *at least* all the deterministic elements in the actual data-generating process, the distribution of γ is nonnormal under the null of a unit root. The distribution itself varies with the set of parameters included in the estimating equation.
2. If the estimated regression includes deterministic regressors that are not in the actual data-generating process, the power of the unit root test against a stationary alternative decreases as additional deterministic regressors are added.
3. If the estimated regression omits an important deterministic trending variable present in the true data-generating process, such as the expression $a_2 t$ in (4.35), the power of the t-statistic test goes to zero as the sample size increases. If the estimated regression omits a nontrending variable (i.e., the mean or a change in the mean), the t-statistic is consistent, but the finite sample power is adversely affected and decreases as the magnitude of the coefficient on the omitted component increases.

4. Estimating (4.13) or (4.14), we observe that the τ_μ, τ_τ, ϕ_1, ϕ_2, and ϕ_3 statistics have the *asymptotic* distributions tabulated by Dickey and Fuller (1979, 1981). The critical values of the various statistics depend on sample size. However, the *sample* variance of $\{y_t\}$ will be dominated by the presence of a trend or drift. We saw an example of this phenomenon in Figure 3.12 of Chapter 3. The time path of the random walk plus drift model in graph (b) is swamped by the presence of the drift term. The fact that the stochastic trend is precisely the same as in graph (a) has little effect on the overall appearance of the series. Although the proof is beyond the scope of this text, the τ_μ and τ_τ statistics converge to the standardized normal. Specifically,

$$\sum_{t=1}^{T} y_t^2 \Rightarrow a_2^2 T^5\ 20 \qquad \textit{if } a_2 \neq 0$$
$$\Rightarrow a_0^2 T^3\ 3 \qquad \textit{if } a_0 \neq 0 \text{ and } a_2 = 0$$

Only when both a_0 and a_2 equal zero in the regression equation and data-generating process do the nonstandard Dickey–Fuller distributions dominate. If the data-generating process is known to contain a trend or drift, *the null hypothesis $\gamma = 0$ can be tested using the standardized normal distribution.*

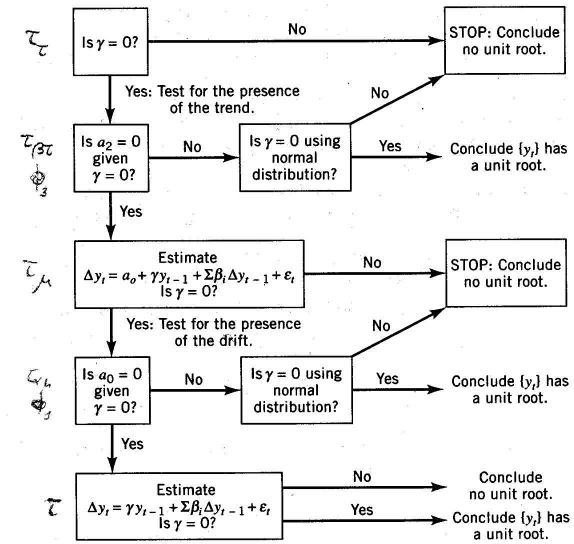
The direct implication of these four findings is that the researcher may fail to reject the null hypothesis of a unit root because of a misspecification concerning the deterministic part of the regression. Too few or too many regressors may cause a failure of the test to reject the null of a unit root. Although we can never be sure that we are including the appropriate deterministic regressors in our econometric model, there are some useful guidelines. Doldado, Jenkinson, and Sosvilla-Rivero (1990) suggest the following procedure to test for a unit root when the form of the data-generating process is unknown. The following is a straightforward modification of their method:

STEP 1: As shown in Figure 4.7, start with the least restrictive of the plausible models (which will generally include a trend and drift) and use the τ_τ statistic to test the null hypothesis $\gamma = 0$. Unit root tests have low power to reject the null hypothesis; hence, if the null hypothesis of a unit root is *rejected,* there is no need to proceed. Conclude that the $\{y_t\}$ sequence does not contain a unit root.

STEP 2: If the null hypothesis is *not rejected,* it is necessary to determine whether too many deterministic regressors were included in Step 1 above.[11] Test for the significance of the trend term under the null of a unit root (e.g., use the $\tau_{\beta\tau}$ statistic to test the significance of a_2). You should try to gain additional confirmation for this result by testing the hypothesis $a_2 = \gamma = 0$ using the ϕ_3 statistic. If the trend is not significant, proceed to Step 3. Otherwise, if the trend is significant, retest for the presence of a unit root (i.e., $\gamma = 0$) using the standardized normal distribution. After all, if a trend is inappropriately included in the estimating equation, the limiting distribution of a_2 is the standardized normal. If the null of a unit root is rejected, proceed no further; conclude that the $\{y_t\}$ sequence does not contain a unit root. Otherwise, conclude that the $\{y_t\}$ sequence contains a unit root.

STEP 3: Estimate (4.35) without the trend [i.e., estimate a model in the form of (4.13)]. Test for the presence of a unit root using the τ_μ statistic. If the null is rejected, conclude that the model does not contain a unit root. If the null hypothesis of a unit root is not rejected, test for the significance of the constant (e.g., use the $\tau_{\alpha\mu}$ statistic to test the significance of a_0 given $\gamma = 0$). Additional confirmation of this result can be obtained by testing the hypothesis $a_0 = \gamma = 0$ using the ϕ_1 statistic. If the drift is not significant, estimate an equation in the form of (4.12) and proceed to Step 4. If the drift is significant, test for the presence of a unit root using the standardized

Figure 4.7 A procedure to test for unit roots.
Estimate $\Delta y_t = a_o + \gamma y_{t-1} + a_2 t + \Sigma\beta_i \Delta y_{t-i} + \varepsilon_t$



normal. If the null hypothesis of a unit root is rejected, conclude that the $\{y_t\}$ sequence does not contain a unit root. Otherwise, conclude that the $\{y_t\}$ sequence contains a unit root.

STEP 4: Estimate (4.35) without the trend or drift, that is, estimate a model in the form of (4.12). Use τ to test for the presence of a unit root. If the null hypothesis of a unit root is rejected, conclude that the $\{y_t\}$ sequence does not contain a unit root. Otherwise, conclude that the $\{y_t\}$ sequence contains a unit root.

Remember, this procedure is not designed to be applied in a completely mechanical fashion. Plotting the data is usually an important indicator of the presence of deterministic regressors. The data shown in Figure 4.1 could hardly be said to contain a deterministic trend. Moreover, theoretical considerations might suggest the appropriate regressors. The efficient market hypothesis is inconsistent with the presence of a deterministic trend in an asset market price. However, the procedure is a sensible way to test for unit roots when the form of the data-generating process is completely unknown.

GDP and Unit Roots

Although the methodology outlined in Figure 4.7 can be very useful, it does have its problems. Each step in the procedure involves a test that is conditioned on all the previous tests being correct; the significance level of each of the cascading tests is impossible to ascertain.

The procedure and its inherent dangers are nicely illustrated by trying to determine if real gross domestic product (GDP) has a unit root. The data are contained in the file entitled US.WK1 on the data disk; it is a good idea to replicate the results reported below. If we use quarterly data over the 1960:1 to 1991:4 period, the correlogram of the logarithm of real GDP exhibits slow decay. However, the first 12 autocorrelations and partial autocorrelations of the logarithmic first difference are

ACF of the logarithmic first difference of real GDP:

Lag 1:	0.3093189	0.2316683	0.0572363	0.0556556	−0.0604932	0.0336679
7:	−0.0476200	−0.1453376	−0.0461222	0.0600729	0.0101171	−0.1695323

PACF of the logarithmic first difference of real GDP:

Lag 1:	0.3093189	0.1503780	−0.0567524	0.0220048	−0.0876589	0.0696282
7:	−0.0507211	−0.1605942	0.0669240	0.1200468	−0.0353431	−0.2423071

Despite the somewhat large partial correlation at lag 12, the Box–Jenkins procedure yields the ARIMA(0, 1, 2) model:

$$\Delta LGDP_t = \underset{(0.001144)}{0.007018} + (1 + \underset{(0.088250)}{0.262169}L + \underset{(0.082663)}{0.197547}L^2)\epsilon_t$$

where $LGDP_t = \log(GDP_t)$, so that $\Delta LGDP_t$ is the growth rate of real GDP, and standard errors are in parentheses.

The model is well estimated in that the residuals appear to be white-noise and all coefficients are of high quality. For our purposes, the interesting point is that the $\Delta\log(GDP_t)$ series appears to be a stationary process. Integrating suggests that $\log(GDP_t)$ has a stochastic and deterministic trend. The deterministic quarterly growth rate of 0.007018—close to a 3% annual rate—appears to be quite reasonable. Now consider the three augmented Dickey–Fuller equations with t-statistics in parentheses:

$$\Delta LGDP_t = \underset{(2.56548)}{0.79018} - \underset{(-2.54309)}{0.05409}LGDP_{t-1} + \underset{(2.27941)}{0.000348}t + \underset{(2.83349)}{0.24961}\Delta LGDP_{t-1} + \underset{(1.94841)}{0.17273}\Delta LGDP_{t-2} \quad (4.37)$$
$$RSS = 0.0089460783$$

$$\Delta LGDP_t = \underset{(2.05219)}{0.09600} - \underset{(-1.96196)}{0.00611}LGDP_{t-1} + \underset{(2.64113)}{0.23613}\Delta LGDP_{t-1} + \underset{(1.52736)}{0.13535}\Delta LGDP_{t-2} \quad (4.38)$$
$$RSS = 0.0093334206$$

$$\Delta LGDP_t = \underset{(3.82135)}{0.000279}LGDP_{t-1} + \underset{(2.93959)}{0.26331}\Delta LGDP_{t-1} + \underset{(1.72964)}{0.15443}\Delta LGDP_{t-2} \quad (4.39)$$
$$RSS = 0.0096582756$$

From (4.37), the t-statistic for the null hypothesis $\gamma = 0$ is −2.54309. Critical values with 125 usable observations are not reported in the Dickey–Fuller table.[12] However, with 100 observations, the critical value of τ_τ at the 5% significance level is −3.45; hence, it is not possible to reject the null hypothesis of a unit root given the presence of the drift term and time trend.

The power of the test may have been reduced due to the presence of an unnecessary time trend and/or drift term. In Step 2, you test for the presence of the time trend given the presence of a unit root. In (4.37), the t-statistic for the null hypothesis that $a_2 = 0$ is 2.27941. Do not let this large value fool you into thinking that a_2 is significantly different from zero. Remember, in the presence of a unit root, you cannot use the critical values of a t-table; instead, the appropriate critical values are given by the Dickey–Fuller $\tau_{\beta\tau}$ statistic. As you can see in Table 4.1, the critical value of $\tau_{\beta\tau}$ at the 5% significance level is 2.79; hence, it is reasonable to conclude that $a_2 = 0$. The ϕ_3 statistic to test the joint hypothesis $a_2 = \gamma = 0$ reconfirms this result. If we view (4.37) as the unrestricted model and (4.39) as the restricted model, there are two restrictions and 120 degrees of freedom in the unrestricted model; the ϕ_3 statistic is

$$\phi_3 = [(0.0096582756 - 0.0089460783)/2]/(0.0089460783/120)$$
$$= 4.7766$$

Since the critical value of ϕ_3 is 6.49, it is possible to conclude that the restriction $a_2 = \gamma = 0$ is not binding. Thus, proceed to Step 3 where you estimate the model without the trend. In (4.38), the t-statistic for the null hypothesis $\gamma = 0$ is -1.96196. Since the critical value of the τ_μ statistic is -2.89 at the 5% significance level, the null hypothesis of a unit root is not rejected at conventional significance levels. Again, the power of this test will have been reduced if the drift term does not belong in the model. To test for the presence of the drift, use the $\tau_{\alpha\tau}$ statistic. The calculated t-statistic is 2.05219, whereas the critical value at the 5% significance level is 2.54. The ϕ_1 statistic also suggests that the drift term is zero. Comparing (4.38) and (4.39), we obtain

$$\phi_1 = (0.0096582756 - 0.0093334206)/(0.0093334206/121)$$
$$= 4.21147365$$

Proceeding to Step 4 yields (4.39). The point is that the procedure has worked itself into an uncomfortable corner. The problem is that the positive coefficient for γ (i.e., the estimated value of $\gamma = 0.000279$ is almost four standard deviations from zero) suggests an *explosive* process. In Step 3, it was probably unwise to conclude that the drift term is equal to zero. As you should verify in Exercise 4 at the end of this chapter, the simple Box–Jenkins ARIMA(0, 1, 2) model with an intercept of 0.007018 performs better than any of the alternatives.

STATISTICAL TABLES

Table A **Empirical Cumulative Distribution of τ**

Probability of a Smaller Value **Sample Size**	**0.01**	**0.025**	**0.05**	**0.10**	**0.90**	**0.95**	**0.975**	**0.99**
No Constant or Time ($a_0 = a_2 = 0$)				τ				
25	−2.66	−2.26	−1.95	−1.60	0.92	1.33	1.70	2.16
50	−2.62	−2.25	−1.95	−1.61	0.91	1.31	1.66	2.08
100	−2.60	−2.24	−1.95	−1.61	0.90	1.29	1.64	2.03
250	−2.58	−2.23	−1.95	−1.62	0.89	1.29	1.63	2.01
300	−2.58	−2.23	−1.95	−1.62	0.89	1.28	1.62	2.00
∞	−2.58	−2.23	−1.95	−1.62	0.89	1.28	1.62	2.00
Constant ($a_2 = 0$)				τ_τ				
25	−3.75	−3.33	−3.00	−2.62	−0.37	0.00	0.34	0.72
50	−3.58	−3.22	−2.93	−2.60	−0.40	−0.03	0.29	0.66
100	−3.51	−3.17	−2.89	−2.58	−0.42	−0.05	0.26	0.63
250	−3.46	−3.14	−2.88	−2.57	−0.42	−0.06	0.24	0.62
500	−3.44	−3.13	−2.87	−2.57	−0.43	−0.07	−0.24	0.61
∞	−3.43	−3.12	−2.86	−2.57	−0.44	−0.07	0.23	0.60
Constant + time				τ_τ				
25	−4.38	−3.95	−3.60	−3.24	−1.14	−0.80	−0.50	−0.15
50	−4.15	−3.80	−3.50	−3.18	−1.19	−0.87	−0.58	−0.24
100	−4.04	−3.73	−3.45	−3.15	−1.22	−0.90	−0.62	−0.28
250	−3.99	−3.69	−3.43	−3.13	−1.23	−0.92	−0.64	−0.31
500	−3.98	−3.68	−3.42	−3.13	−1.24	−0.93	−0.65	−0.32
∞	−3.96	−3.66	−3.41	−3.12	−1.25	−0.94	−0.66	−0.33

Source: This table was constructed by David A. Dickey using Monte Carlo methods. Standard errors of the estimates vary, but most are less than 0.20. The table is reproduced from Wayne Fuller. *Introduction to Statistical Time Series.* (New York: John Wiley). 1976.

Response Surface Study, MacKinnon 1991

Dickey-Fuller unit-root critical values

$c_\alpha = \beta_0 + \beta_1 T^{-1} + \beta_2 T^{-2}$

	No constant or trend			Constant but no trend			Constant and trend		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
β_0	−2.5658	1.9393	−1.6156	−3.4335	−2.8621	−2.5671	−3.9638	−3.4126	−3.1279
β_1	−1.96	−0.398	−0.181	−5.999	−2.738	−1.438	−8.353	−4.039	−2.418
β_2	−10.04	0	0	−29.25	−8.36	−4.48	−47.44	−17.83	−7.58

T									
25	−2.66	−1.96	−1.62	−3.72	−2.98	−2.63	−4.37	−3.60	−3.24
50	−2.61	−1.95	−1.62	−3.57	−2.92	−2.60	−4.15	−3.50	−3.18
75	−2.59	−1.94	−1.62	−3.52	−2.90	−2.59	−4.08	−3.47	−3.16
100	−2.59	−1.94	−1.62	−3.50	−2.89	−2.58	−4.05	−3.45	−3.15
200	−2.58	−1.94	−1.62	−3.46	−2.88	−2.57	−4.01	−3.43	−3.14
∞	−2.57	−1.94	−1.62	−3.43	−2.86	−2.57	−4.96	−3.41	−3.13