Chapter 12 Regression with Time-Series Data: Nonstationary Variables

Chapter Contents

Regression with Time-Series Data: Nonstationary Variables

- The aim is to describe how to estimate regression models involving nonstationary variables
  - The first step is to examine the time-series concepts of **stationarity** (and **nonstationarity**) and how we distinguish between them.
  - **Cointegration** is another important related concept that has a bearing on our choice of a regression model

12.1 Stationary and Nonstationary Variables 1 of 3

- The <u>change in a variable</u> is an important concept
  - The change in a variable y_t, also known as its **first difference**, is given by $\Delta y_t = y_t - y_{t-1}$.
  - Δy_t is the change in the value of the variable y from period t - 1 to period t
- Observe how the GDP variable displays upward trending behavior, while the other series "wander up and down"

FIGURE 12.1 U.S. Economic Time Series 1 of 2

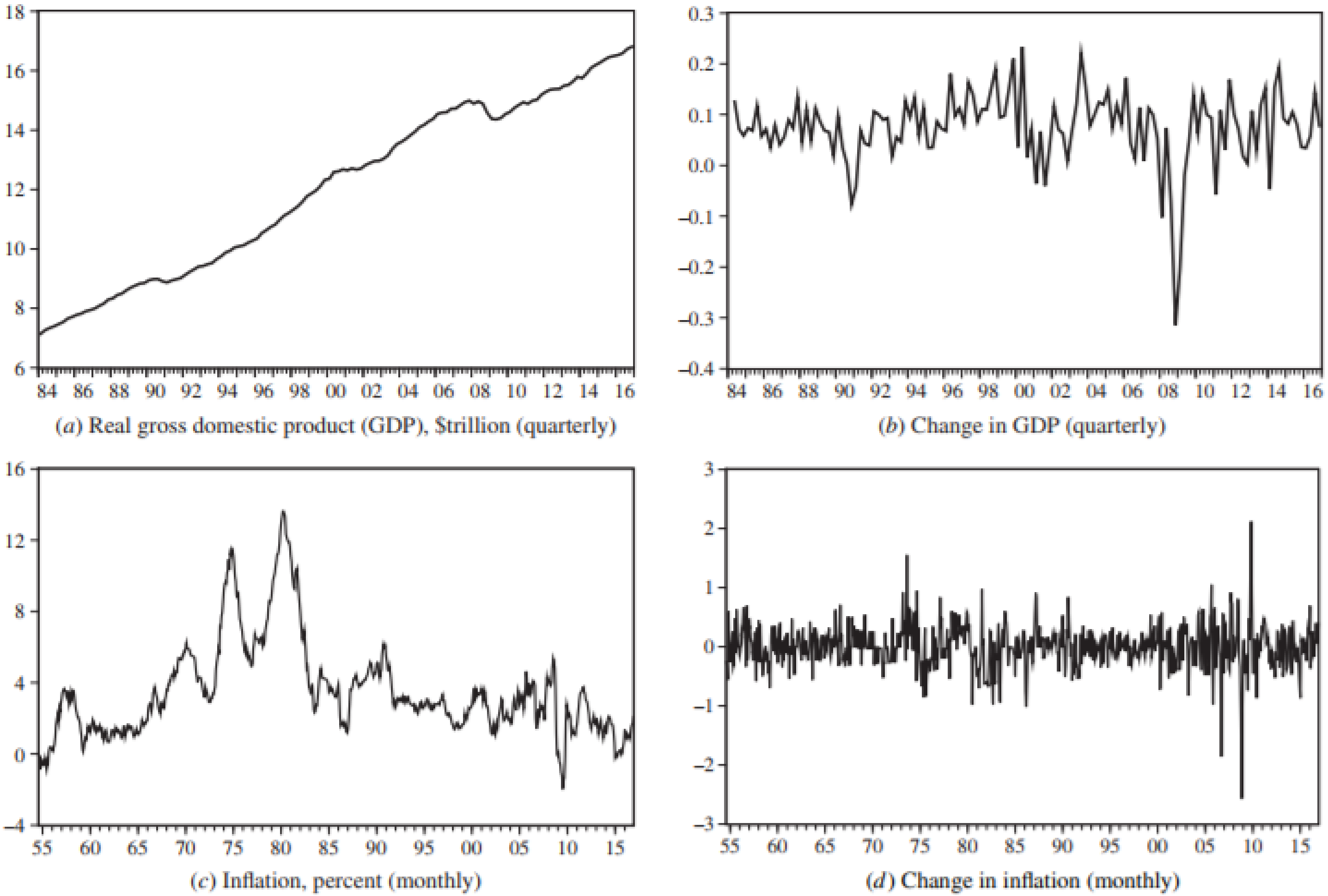
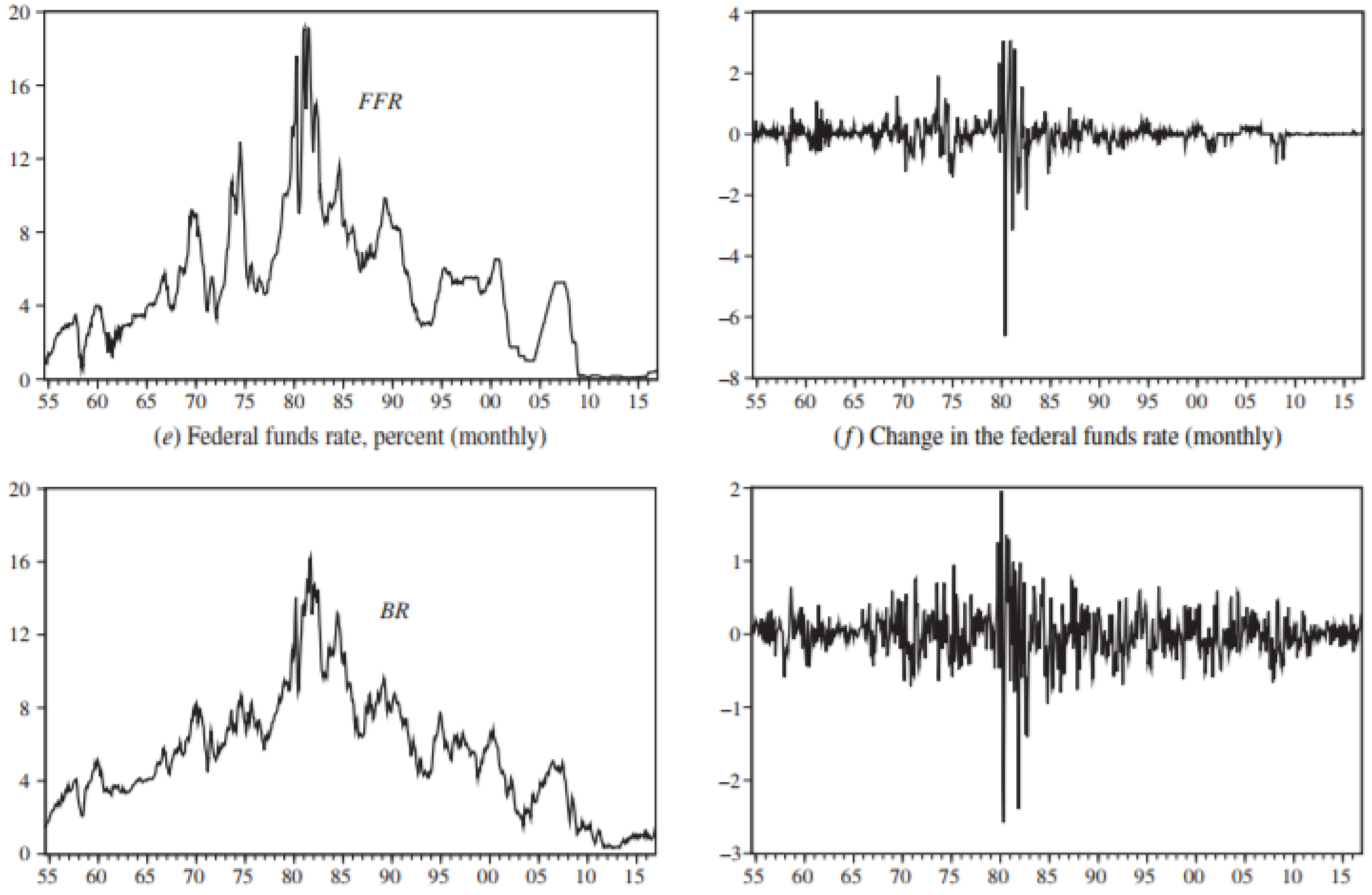


FIGURE 12.1 U.S. Economic Time Series 2 of 2



(e) Federal funds rate, percent (monthly)

(f) Change in the federal funds rate (monthly)

(g) Three-year bond rate, percent (monthly)

(h) Change in the bond rate (monthly)

12.1 Stationary and Nonstationary Variables 2 of 3

- Recall that a stationary time series y_t has mean and variance that are <u>constant</u> over time
  - and that the covariance between two values from the series depends only on the <u>length of time</u> separating the two values
  - and not on the actual times at which the variables are observed

12.1 Stationary and Nonstationary Variables 3 of 3

- That is:
- (12.1a) $E(y_t) = \mu$ (constant mean)
- (12.1b) $\text{var}(y_t) = \sigma^2$ (constant variance)
- (12.1c) $\text{cov}(y_t, y_{t+s}) = \text{cov}(y_t, y_{t-s}) = \gamma_s$ (covariance depends on *s*, not *t*)
- Another characteristic of nonstationary variables is that their sample

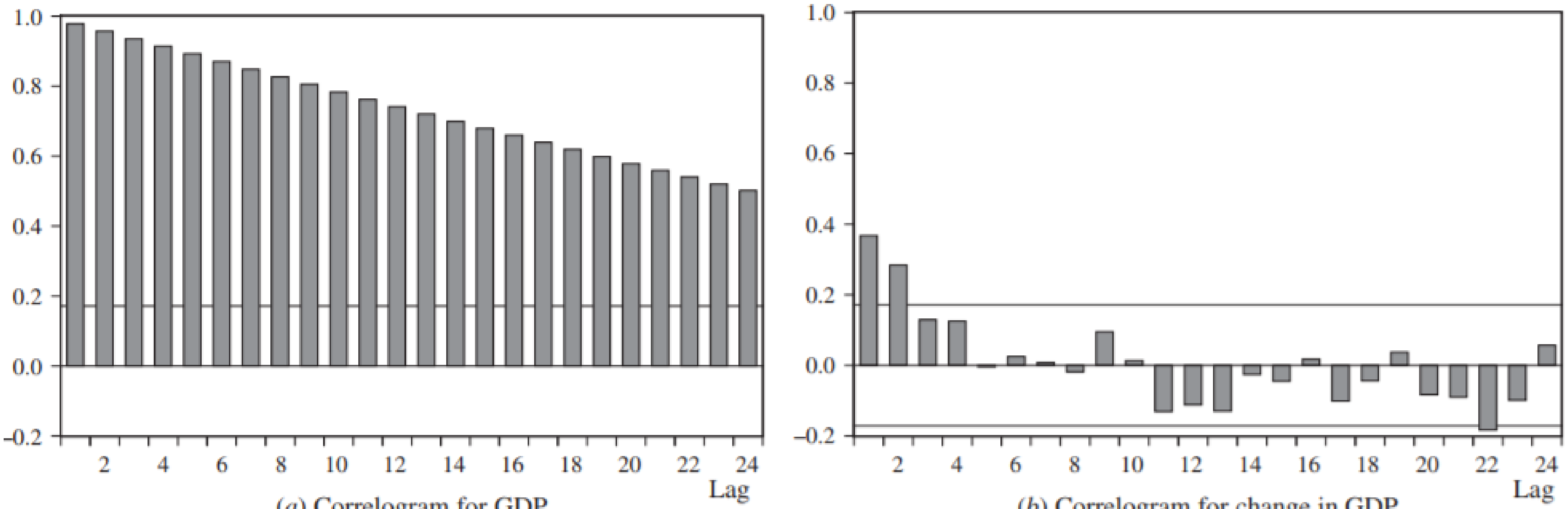
  autocorrelations remain large at long lags
- The sample autocorrelations of nonstationary series exhibit **strong dependence**

Table 12.1 Sample Means of Time Series Shown in Figure 12.1

TABLE 12.1 **Sample Means of Time Series Shown in Figure 12.1**

Variable		Sample periods	
	GDP	1948Q2 to 2000Q3	2000Q4 to 2016Q4
	Other variables	1954M8 to 1985M10	1985M11 to 2016M12
Real GDP (a)		9.56	14.68
Inflation rate (c)		4.42	2.59
Federal funds rate (e)		6.20	3.65
Bond rate (g)		6.56	4.29
Change in GDP (b)		0.083	0.065
Change in the inflation rate (d)		0.01	−0.003
Change in the federal funds rate (f)		0.02	−0.02
Change in the bond rate (h)		0.02	−0.02

FIGURE 12.2 Correlograms for GDP and the change in GDP



(*a*) Correlogram for GDP (*b*) Correlogram for change in GDP

12.1.1 Trend Stationary Variables 1 of 3

- Nonstationary variables that wander up and down, trending in one direction and then the other, are said to possess a stochastic trend
- Definite trends, upward or downward, can be attributable to a stochastic trend or a deterministic trend
- Variables that are stationary after "subtracting out" a deterministic trend are called trend stationary
- Consider the model:

$$(12.2)\ y_t = c_1 + c_2 t + u_t$$

12.1.1 Trend Stationary Variables 2 of 3

- If we focus just on the trend and assume any change in the error is zero then the coefficient c_2 gives the change in y from a one period to the next
- Since fluctuations are given by changes in the error term
- (12.3) $u_t = y_t - (c_1 + c_2 t)$
- y_t is trend stationary if u_t is stationary
- use least squares to find estimates $\hat{c}_1$ and $\hat{c}_2$
- (12.4) $\hat{u}_t = y_t - (\hat{c}_1 + \hat{c}_2 t)$

12.1.1 Trend Stationary Variables 3 of 3

- Another popular trend is one where, on average, a variable is growing at a constant percentage rate
- A model with this property, with an error term included, is
- (12.9) $\ln(y_t) = a_1 - a_2 t + u_t$
- In this case
  - the deterministic trend for y_t is $\exp(a_1 - a_2 t)$, and $\ln(y_t)$ will be trend stationary if u_t is stationary

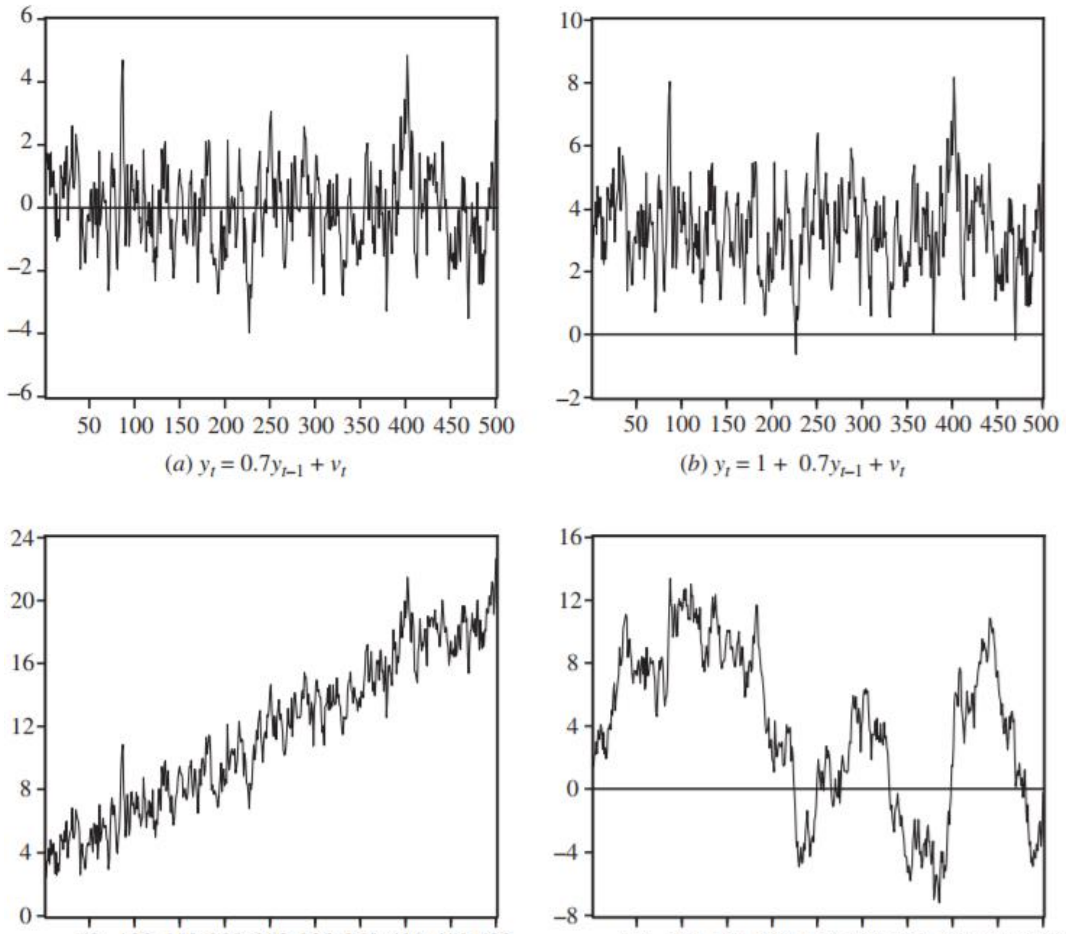
12.1.2 The First-order Autoregressive Model 1 of 5

- To develop a framework for modeling nonstationary variables that possess a stochastic trend, we begin by revising the first-order autoregressive **AR(1)** model
- The econometric model generating a time-series variable y_t is called a **stochastic** or **random process**
- A sample of observed y_t values is called a particular **realization** of the stochastic process
  - It is one of many possible paths that the stochastic process could have taken

12.1.2 The First-order Autoregressive Model 2 of 5

- The autoregressive model of order one, the AR(1) model, is a useful univariate time series model for explaining the difference between stationary and nonstationary series:
- (12.12) $y_t = \rho y_{t-1} + v_t, \qquad |\rho| < 1$
  - The errors v_t are independent, with zero mean and constant variance σ_v^2, and may be normally distributed
  - The errors are sometimes known as "shocks" or "innovations"

FIGURE 12.4 Time-series models 1 of 2



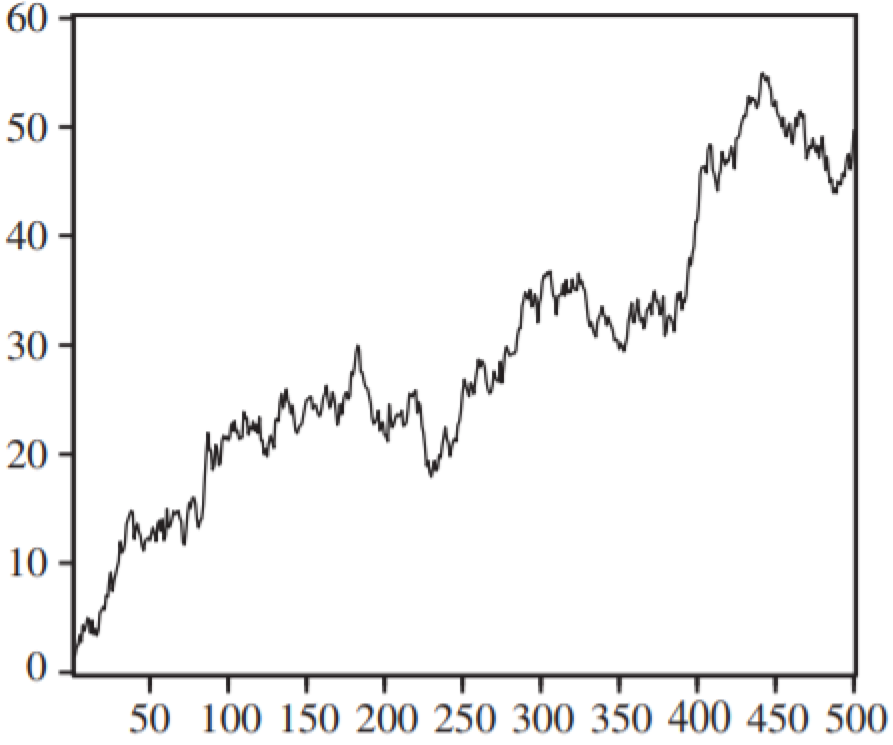
(a) $y_t = 0.7y_{t-1} + v_t$

(b) $y_t = 1 + 0.7y_{t-1} + v_t$

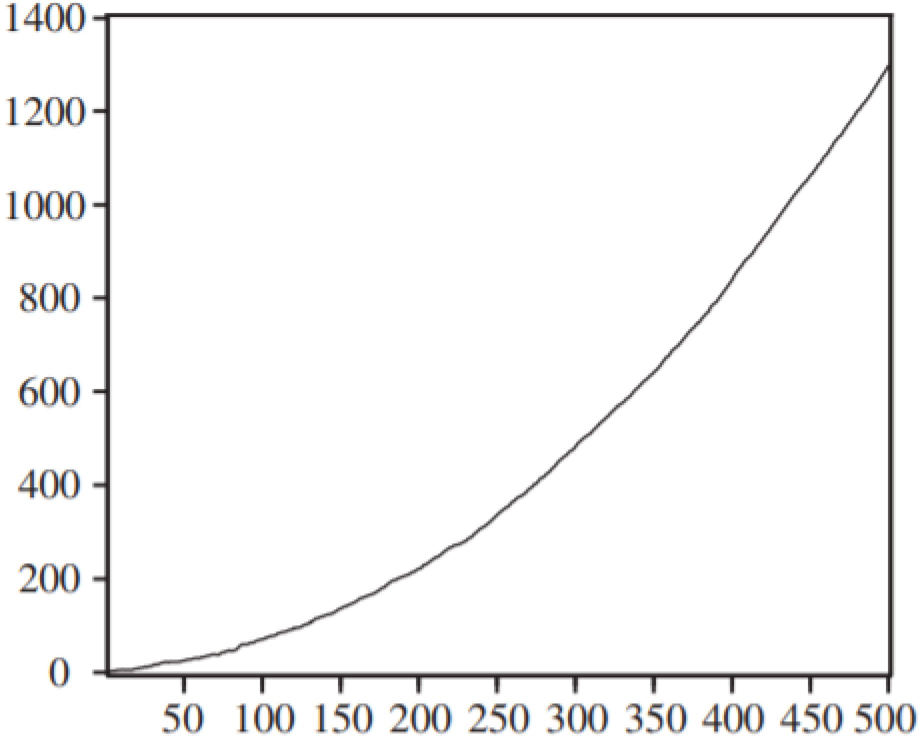
(c) $y_t = 1 + 0.01t + 0.7y_{t-1} + v_t$

(d) $y_t = y_{t-1} + v_t$

FIGURE 12.4 Time-series models 2 of 2



(e) $y_t = 0.1 + y_{t-1} + v_t$



(f) $y_t = 0.1 + 0.01t + y_{t-1} + v_t$

12.1.2 The First-order Autoregressive Model 3 of 5

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

$$y_1 = \rho y_0 + v_1$$

$$y_2 = \rho y_1 + v_2 = \rho(\rho y_0 + v_1) + v_2 = \rho^2 y_0 + \rho v_1 + v_2$$

$$\vdots$$

$$y_t = v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \ldots.. + \rho^t y_0$$

12.1.2 The First-order Autoregressive Model 4 of 5

- The mean of y_t is:

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

- Real-world data rarely have a zero mean
  - We can introduce a nonzero mean μ as:

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

  - Or
  - (12.13) $\quad y_t = \alpha + \rho y_{t-1} + v_t, \qquad |\rho| < 1$

12.1.2 The First-order Autoregressive Model 5 of 5

- An example of a time series that follows this model, with $\alpha = 1$ and $\rho = 0.7$:

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

- An extension to (12.12) is to consider an AR(1) model fluctuating <u>around a linear trend</u>: $(\mu + \delta t)$
  - Let the "de-trended" series (y_t -μ - δt) behave like an autoregressive model:

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

  - Or

(12.4) $y_t = \alpha + \rho y_{t-1} + \lambda t + v_t$

12.1.3 Random Walk Models 1 of 6

- Consider the special case of $\rho = 1$:
- (12.15) $y_t = y_{t-1} + v_t$
- This model is known as the random walk model
  - These time series are called **random walks** because they appear to wander slowly upward or downward with no real pattern
  - the values of sample means calculated from subsamples of observations will be dependent on the sample period
  - This is a characteristic of nonstationary series

12.1.3 Random Walk Models 2 of 6

- We can understand the "wandering" by recursive substitution:

$$y_1 = y_0 + v_1$$

$$y_2 = y_1 + v_2 = (y_0 + v_1) + v_2 = y_0 + \sum_{s=1}^{2} v_s$$

$$\vdots$$

$$y_t = y_{t-1} + v_t = y_0 + \sum_{s=1}^{t} v_s$$

- The random walk model contains an initial value y_0

12.1.3 Random Walk Models 3 of 6

- A The term $\sum_{s=1}^{t} v_s$ is often called the **stochastic trend**
  - This term arises because a stochastic component v_t is added for each time t, and because it causes the time series to trend in unpredictable directions
- Recognizing that the v_t are independent, taking the expectation and the variance of y_t yields, for a fixed initial value y_0: $E(y_t) = y_0 + E(v_1 + v_2 + \ldots + v_t) = y_0$

  $$\operatorname{var}(y_t) = \operatorname{var}(v_1 + v_2 + \ldots + v_t) = t\sigma_v^2$$

  - The random walk has a mean equal to its initial value and a variance that increases over time, eventually becoming infinite

12.1.3 Random Walk Models 4 of 6

- Another nonstationary model is obtained by adding a constant term:
- (12.16) $y_t = \delta + y_{t-1} + v_t$
  - This model is known as the **random walk with drift**
- A better understanding is obtained by applying recursive substitution:

$$y_1 = \delta + y_0 + v_1$$

$$y_2 = \delta + y_1 + v_2 = \delta + (\delta + y_0 + v_1) + v_2 = 2\delta + y_0 + \sum_{s=1}^{2} v_s$$

$$\vdots$$

$$y_t = \delta + y_{t-1} + v_t = t\delta + y_0 + \sum_{s=1}^{t} v_s$$

12.1.3 Random Walk Models 5 of 6

- The term $t\delta$ a **deterministic trend** component
  - It is called a deterministic trend because a fixed value δ is added for each time t
  - The variable y wanders up and down as well as increases by a fixed amount at each time t
  - The mean and variance of y_t are:

$$E(y_t) = t\delta + y_0 + E(v_1 + v_2 + v_3 + \ldots + v_t) = t\delta + y_0$$

$$\text{var}(y_t) = \text{var}(v_1 + v_2 + v_3 + \ldots + v_t) = t\sigma_v^2$$

12.1.3 Random Walk Models 6 of 6

- We can extend the random walk model even further by adding a <u>time trend</u>:
- (12.17) $y_t = \alpha + \delta t + y_{t-1} + v_t$
- The addition of a time-trend variable *t* strengthens the trend behavior:

$$y_1 = \alpha + \delta + y_0 + v_1$$

$$y_2 = \alpha + \delta 2 + y_1 + v_2 = \alpha + 2\delta + (\alpha + \delta + y_0 + v_1) + v_2 = 2\alpha + 3\delta + y_0 + \sum_{s=1}^{2} v_s$$

$$\vdots$$

$$y_t = \alpha + \delta t + y_{t-1} + v_t = t\alpha + \left(\frac{t(t+1)}{2}\right)\delta + y_0 + \sum_{s=1}^{t} v_s$$

- where we used: $1 + 2 + 3 + \cdots + t = t(t+1)/2$

12.2 Consequences of Stochastic Trends

- Now we consider the implications of estimating regressions involving variables with stochastic trends
- A consequence of proceeding with the regression involving nonstationary variables with stochastic trends is that OLS estimates no longer have approximate normal distributions in large samples.
- One particular hazard is that two totally independent random walks can appear to have a strong linear relationship when none exists
  - Outcomes of this nature have been given the name spurious regressions

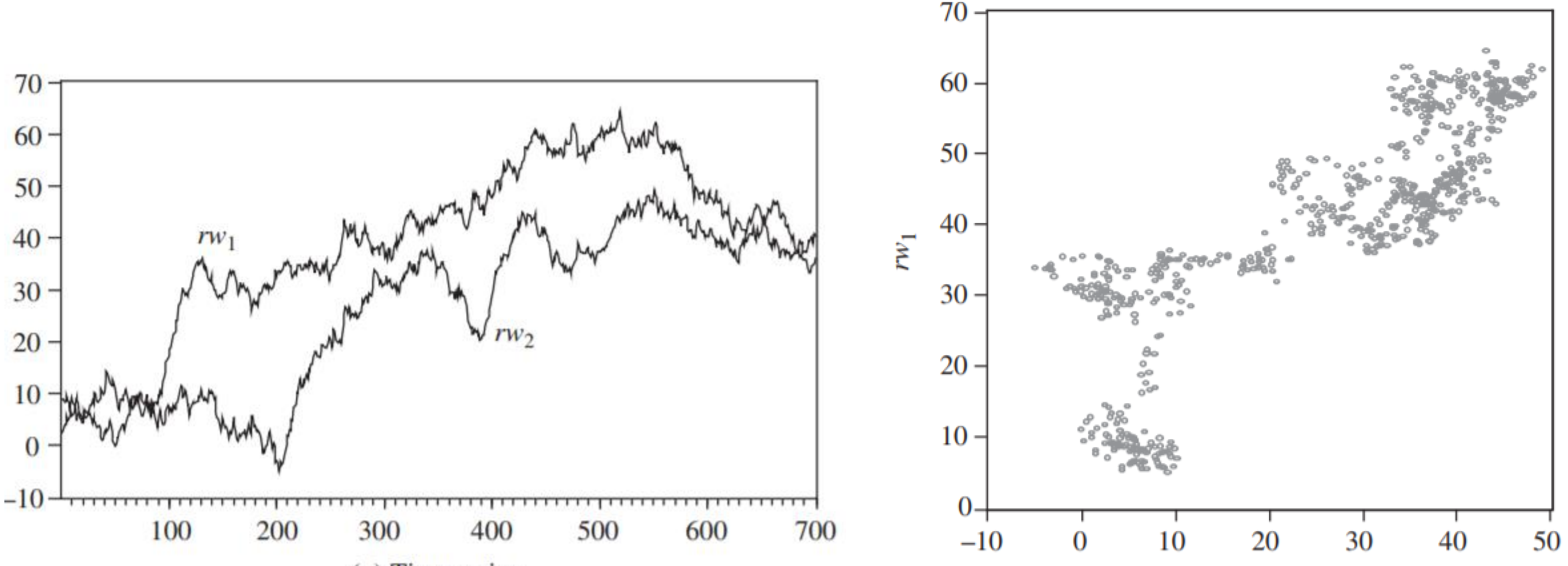
EXAMPLE 12.3 A Regression With Two Random Walks 1 of 3

- Consider two independent random walks:

$$rw_1:\ \ y_t = y_{t-1} + v_{1t}$$
$$rw_2:\ \ x_t = x_{t-1} + v_{2t}$$

- where v_{1t} and v_{2t} are independent N(0, 1) random error
- These series were generated independently and, in truth, have no relation to one another
- Yet when plotted we see a positive relationship between them

FIGURE 12.5 Time series and scatter plot of two random walk variables



(a) Time series

(b) Scatter plot

EXAMPLE 12.3 A Regression With Two Random Walks 2 of 3

- A simple regression of series one (rw_1) on series two (rw_2) yields:

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

- And that the estimated slope is significantly different from zero
  - the t-statistic is huge
- These results are completely meaningless, or spurious
  - The apparent significance of the relationship is false

EXAMPLE 12.3 A Regression With Two Random Walks 3 of 3

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

12.3 Unit Root Tests for Stationarity

- There are many tests for determining whether a series is stationary or nonstationary
- The most popular is the **Dickey–Fuller** test

12.3.1 Unit Roots 1 of 2

- The AR(1) process $y_t = \rho y_{t-1} + v_t$ is stationary when $|\rho| < 1$
  - But, when $\rho = 1$, it becomes the nonstationary random walk process
- Consider the more general AR(p) model: $y_t = \alpha + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \cdots + \theta_p y_{t-p} + v_t$ In this model, y_t is stationary if the roots of the polynomial equation
  - (12.18) $\varphi(z) = 1 - \theta_1 z - \theta_1 z^2 - \cdots - \theta_p z^p$
  - are greater than one in absolute value

12.3.1 Unit Roots 2 of 2

- The roots are the values of z that satisfy the equation φ(z) = 0
- The condition for stationarity is |z| > 1, which is the same as $|\theta_1| < 1$
- If, in (12.18), one of the roots is equal to one, then y_t is said to have a unit root
- In higher-order AR models, the conditions for a unit root and for stationarity, written in terms of the parameters $\theta_1, \theta_2, \ldots, \theta_p$, are more complicated

12.3.2 Dickey–Fuller Tests

- There are three variations of the Dickey–Fuller test

1. The alternative hypothesis is that y_t is stationary around a nonzero mean
   - The test equation includes an intercept but no trend term
2. The alternative hypothesis is that y_t is stationary around a linear deterministic trend
   - the test equation includes both intercept and trend terms
3. The alternative hypothesis is that y_t is stationary around a zero mean
   - Both intercept and trend are excluded from the test equation in this case

12.3.3 Dickey–Fuller Test with Intercept and No Trend 1 of 4

- The nonstationary random walk is set up as the null hypothesis
- (12.19) $H_0: y_t = y_{t-1} + v_t$
- The stationary AR(1) process becomes the alternative hypothesis
- (12.20) $H_1: y_t = \alpha + \rho y_{t-1} + v_t \quad |\rho| < 1$
- testing H_0:ρ = 1 against the alternative H_1 :|ρ| < 1, or simply H_1 :ρ < 1
- This one-sided (left tail) test is put into a more convenient form by subtracting y_{t-1} from both sides of (12.20)

12.3.3 Dickey–Fuller Test with Intercept and No Trend 2 of 4

- (12.21): $y_t - y_{t-1} = \alpha + \rho y_{t-1} - y_{t-1} + v_t$

$$\Delta y_t = \alpha + (\rho - 1)y_{t-1} + v_t$$
$$= \alpha + \gamma y_{t-1} + v_t$$

- The hypotheses are:

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

- (12.22) $H_1\!: \rho < 1 \quad \Leftrightarrow \quad H_1\!: \gamma < 0$

- Rejection of the null hypothesis that $\gamma = 0$ implies the series is stationary

- A failure to reject H_0 suggests the series could be nonstationary

12.3.3 Dickey–Fuller Test with Intercept and No Trend 3 of 4

- To test the hypothesis in (12.22), we estimate the test equation (12.21) by OLS
- We use a **τ** (tau) statistic, and its value must be compared to specially generated critical values
- We reject $H_0 : \gamma = 0$ if $\tau \leq \tau_c$
- To test for higher-order AR process
- (12.23) $\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{s=1}^{p-1} a_s \Delta y_{t-s} + v_t$

12.3.3 Dickey–Fuller Test with Intercept and No Trend 4 of 4

- The test procedure for this case uses (12.23) as the test equation but otherwise proceeds just as before
- The test is referred to as the augmented Dickey–Fuller test
- Sufficient lags should be included to eliminate autocorrelation in the errors
- In practice, we always use the augmented Dickey–Fuller test (rather than the nonaugmented version) to ensure the errors are uncorrelated

12.3.4 Dickey–fuller Test with Intercept and Trend

- The Dickey–Fuller test with intercept and trend is designed to discriminate between these two models
- (12.24) $y_t = \alpha + \rho y_{t-1} + \lambda_t + v_t$ $|\rho| < 1$ is the alternative hypothesis
- (12.25) $y_t = \alpha + y_{t-1} + v_t$ becomes the null hypothesis
- the test equation is obtained by subtracting y_{t-1} from both sides of (12.24) and adding augmentation terms to obtain
- (12.26) $\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{s=1}^{p-1} a_s \Delta y_{t-s} + v_t$

12.3.5 Dickey–fuller Test with No Intercept and No Trend

- The test equation of a random walk is:
- (12.27) $\Delta y_t = \gamma y_{t-1} + \sum_{s=1}^{p-1} a_s \Delta y_{t-s} + v_t$
- We test H_0:γ = 0 against H_1:γ < 0
- Most time series measured in terms of their original levels do not have a zero mean.

  However, their first differences $\Delta y_t = y_t - y_{t-1}$ may turn out to have a zero mean

12.3.6 Order of Integration

- Recall that if y_t follows a random walk, then $\gamma = 0$ and the first difference of y_t becomes:

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

- Series like y_t, which can be made stationary by taking the first difference, are said to be **integrated of order one**, and denoted as **I(1)**
  - Stationary series are said to be integrated of order zero, **I(0)**
- In general, the order of integration of a series is the minimum number of times it must be differenced to make it stationary

12.3.7 Other Unit Root Tests 1 of 2

- The power of the Dickey–Fuller tests is low in the sense that they often cannot distinguish between a highly persistent stationary process and a nonstationary process

- The power of the test also diminishes as deterministic terms constant and trend are included in the test equation

- Here we briefly mention other tests that have been developed with a view to improving the power of the test

12.3.7 Other Unit Root Tests 2 of 2

- The ERS test proposes removing the constant/trend effects from the data and performing the unit root test on the residuals
- The PP test adopts a nonparametric approach that assumes a general autoregressive moving-average structure and uses spectral methods to estimate the standard error of the test correlation
- the KPSS test specifies a null hypothesis that the series is stationary or trend stationary
- NP tests suggest various modifications of the PP and ERS tests

12.4 Cointegration 1 of 3

- As a general rule, nonstationary time-series variables should not be used in regression models to avoid the problem of spurious regression
  - There is an exception to this rule
- There is an important case when $e_t = y_t - \beta_1 - \beta_2 x_t$ is a stationary I(0) process
- In this case y_t and x_t are said to be **cointegrated**
  - Cointegration implies that y_t and x_t share similar stochastic trends, and, since the difference e_t is stationary, they never diverge too far from each other

12.4 Cointegration 2 of 3

- The test for stationarity of the residuals is based on the test equation:
- (12.28) $\Delta\hat{e}_t = \gamma\hat{e}_{t-1} + v_t$
- The regression has no constant term because the mean of the regression residuals is zero.
- We are basing this test upon estimated values of the residuals
- The proper critical values for a test of cointegration are given in Table 12.4

Table12.4 Critical Values for the Cointegration Test

TABLE 12.4 **Critical Values for the Cointegration Test**

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

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

12.4 Cointegration 3 of 3

- There are three sets of critical values
- Which set we use depends on whether the residuals $\hat{e}_t$ are derived from:
- (12.29a) $Equation\,1: \quad \hat{e}_t = y_t - b x_t$
- (12.29b) $Equation\,2: \quad \hat{e}_t = y_t - b_2 x_t - b_1$
- (12.29c) $Equation\,3: \quad \hat{e}_t = y_t - b_2 x_t - b_1 - \hat{\delta} t$

Example 12.8 Are the Federal Funds Rate and Bond Rate Cointegrated 1 of 2

- let us test whether y_t = BR_t and x_t = FFR_t, are cointegrated

- The estimated least-squares regression between these variables is

- (12.30)
$$\begin{array}{llll} \widehat{BR}_t = 1.328 + 0.832F_t, & & R^2 = 0.908 \\ (t) & (85.72) & & (12.30) \end{array}$$

- The unit root test for stationarity in the estimated residuals is:

$$\begin{array}{llll} \widehat{\Delta\hat{e}}_t = -0.0817\hat{e}_{t-1} & + 0.223\Delta\hat{e}_{t-1} & - \; 0.177\Delta\hat{e}_{t-1} \\ (\tau \, and \, t) \; (-5.53) & (6.29) & (-4.90) \end{array}$$

Example 12.8 Are the Federal Funds Rate and Bond Rate Cointegrated 2 of 2

- The null and alternative hypotheses in the test for cointegration are:

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

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

- Similar to the one-tail unit root tests, we reject the null hypothesis of no cointegration if $\tau \leq \tau_c$, and we do not reject the null hypothesis that the series are not cointegrated if $\tau > \tau_c$

12.4.1 The Error Correction Model 1 of 4

- Consider a general model that contains lags of *y* and *x*
- Namely, the autoregressive distributed lag (ARDL) model, except the variables are nonstationary:

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

- For simplicity, we shall consider lags up to order one, but the following analysis holds for any order of lags

12.4.1 The Error Correction Model 2 of 4

- If y and x are cointegrated, it means that there is a long-run relationship between them
  - To derive this exact relationship, we set $y_t = y_{t-1} = y$, $x_t = x_{t-1} = x$ and $v_t = 0$
- Imposing this concept in the ARDL, we obtain: $y(1-\theta_1) = \delta + (\delta_0 + \delta_1)x$
  - This can be rewritten in the form:

$$y = \beta_1 + \beta_2 x$$

12.4.1 The Error Correction Model 3 of 4

- Add the term $-y_{t-1}$ to both sides of the equation:

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

- Add the term $-\delta_0 x_{t-1} + \delta_0 x_{t-1}$:

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

- Manipulating this we get:

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

12.4.1 The Error Correction Model 4 of 4

- Using the definitions β_1 and β_2, we get:
- (12.31) $\Delta y_t = -\alpha(y_{t-1} - \beta_1 - \beta_2 x_{t-1}) + \delta_0 \Delta x_t + v_t$
- This is called an error correction equation
- It allows for an underlying or fundamental link between variables (the long-run relationship)
- It allows for short-run adjustments (i.e. changes) between variables, including adjustments to achieve the cointegrating relationship

12.5 Regression When There is No Cointegration 1 of 2

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

12.5 Regression When There is No Cointegration 2 of 2

- If y_t is nonstationary with a stochastic trend and its first difference $\Delta y_t = y_t - y_{t-1}$ is stationary
  - then y_t is I(1) and first-difference stationary
- A suitable regression involving only stationary variables is:
- (12.33) $\Delta y_t = \theta \Delta y_{t-1} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + e_t$
- If y_t and x_t behave like random walks with drift, then:
- (12.34) $\Delta y_t = \alpha + \theta \Delta y_{t-1} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + e_t$

FIGURE 12.7 Regression with time-series data: nonstationary variables

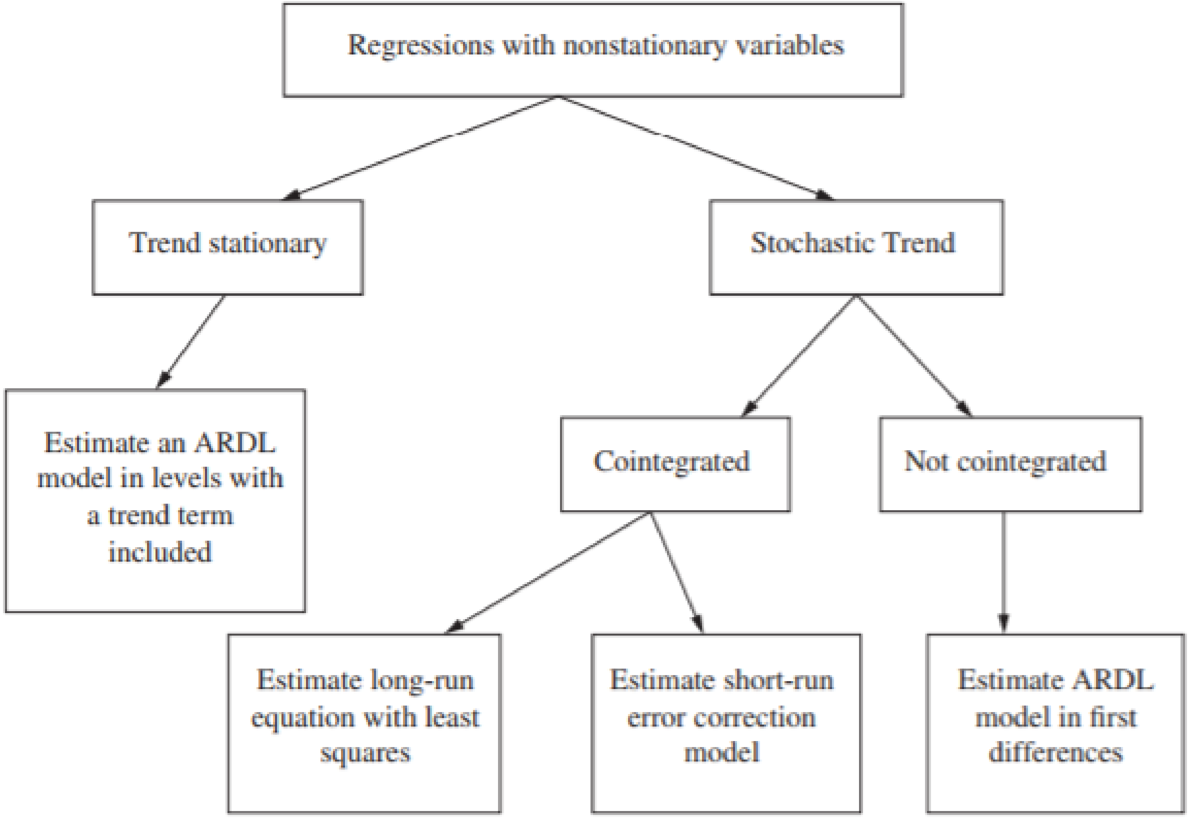


FIGURE 12.7 Regression with time-series data: nonstationary variables.

12.6 Summary

- If variables are stationary, or I(1) and cointegrated, we can estimate a regression relationship between the levels of those variables without fear of encountering a spurious regression.
- If the variables are I(1) and not cointegrated, we need to estimate a relationship in first differences, with or without the constant term
- If they are trend stationary, we can either detrend the series first and then perform regression analysis with the stationary (detrended) variables or, alternatively, estimate a regression relationship that includes a trend variable

Key Words

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

Copyright