Chapter 9
Regression with Time Series Data: Stationary Variables

Walter R. Paczkowski
Rutgers University

Chapter Contents

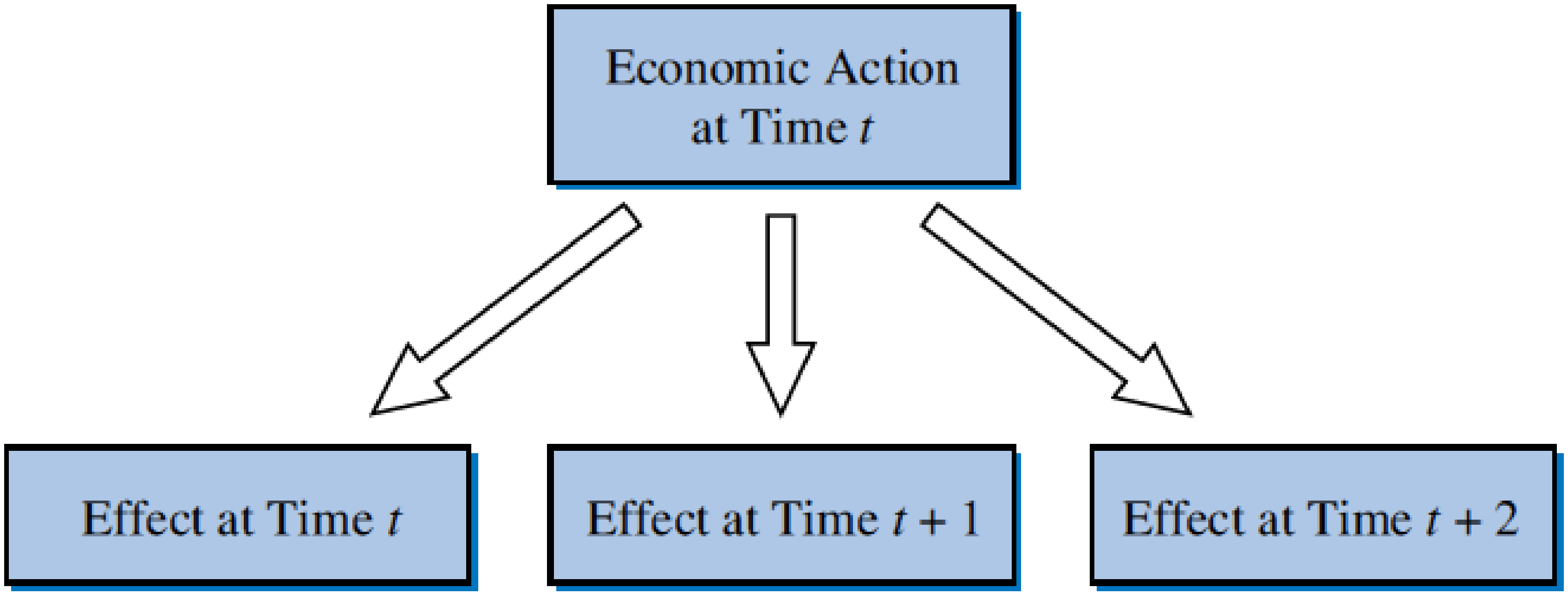
- 9.1 Introduction
- 9.2 Serial Correlation
- 9.3 Other Tests for Serially Correlated Errors
- 9.4 Estimation with Serially Correlated Errors

9.1
Introduction

- When modeling relationships between variables, the nature of the data that have been collected has an important bearing on the appropriate choice of an econometric model
  - Two features of time-series data to consider:
    1. Time-series observations on a given economic unit, observed over a number of time periods, are likely to be correlated
    2. Time-series data have a natural ordering according to time

- There is also the possible existence of dynamic relationships between variables
  - A dynamic relationship is one in which the change in a variable now has an impact on that same variable, or other variables, in one or more future time periods
  - These effects do not occur instantaneously but are spread, or **distributed**, over future time periods

FIGURE 9.1 The distributed lag effect



- Ways to model the dynamic relationship:
  1. Specify that a dependent variable *y* is a function of current and past values of an explanatory variable *x*

Eq. 9.1

$$y_t = f(x_t, x_{t-1}, x_{t-2}, \ldots)$$

- Because of the existence of these lagged effects, Eq. 9.1 is called a distributed lag model

- Ways to model the dynamic relationship (Continued):
  2. Capturing the dynamic characteristics of time-series by specifying a model with a lagged dependent variable as one of the explanatory variables

Eq. 9.2

$$y_t = f(y_{t-1}, x_t)$$

- Or have:

Eq. 9.3

$$y_t = f(y_{t-1}, x_t, x_{t-1}, x_{t-2})$$

– Such models are called **autoregressive distributed lag (ARDL)** models, with "autoregressive" meaning a regression of y_t on its own lag or lags

- Ways to model the dynamic relationship (Continued):

  3. Model the continuing impact of change over several periods via the error term

Eq. 9.4

$$y_t = f(x_t) + e_t \qquad e_t = f(e_{t-1})$$

- In this case e_t is correlated with e_{t-1}
- We say the errors are **serially correlated** or **autocorrelated**

- The primary assumption is Assumption MR4:

$$\text{cov}\left(y_i, y_j\right) = \text{cov}\left(e_i, e_j\right) = 0 \quad \text{for } i \neq j$$

  - For time series, this is written as:

$$\text{cov}\left(y_t, y_s\right) = \text{cov}\left(e_t, e_s\right) = 0 \quad \text{for } t \neq s$$

– The dynamic models in Eqs. 9.2, 9.3 and 9.4 imply correlation between y_t and y_{t-1} or e_t and e_{t-1} or both, so they clearly violate assumption MR4

- A stationary variable is one that is not explosive, nor trending, and nor wandering aimlessly without returning to its mean.
- It has the following properties:
  - $E(y_t) = \mu$
  - $Var(y_t) = \sigma^2$
  - $Cov(y_t, y_{t-s}) = \gamma_s$. The covariance between two different time periods depends on the distance between them.

FIGURE 9.2 (a) Time series of a stationary variable

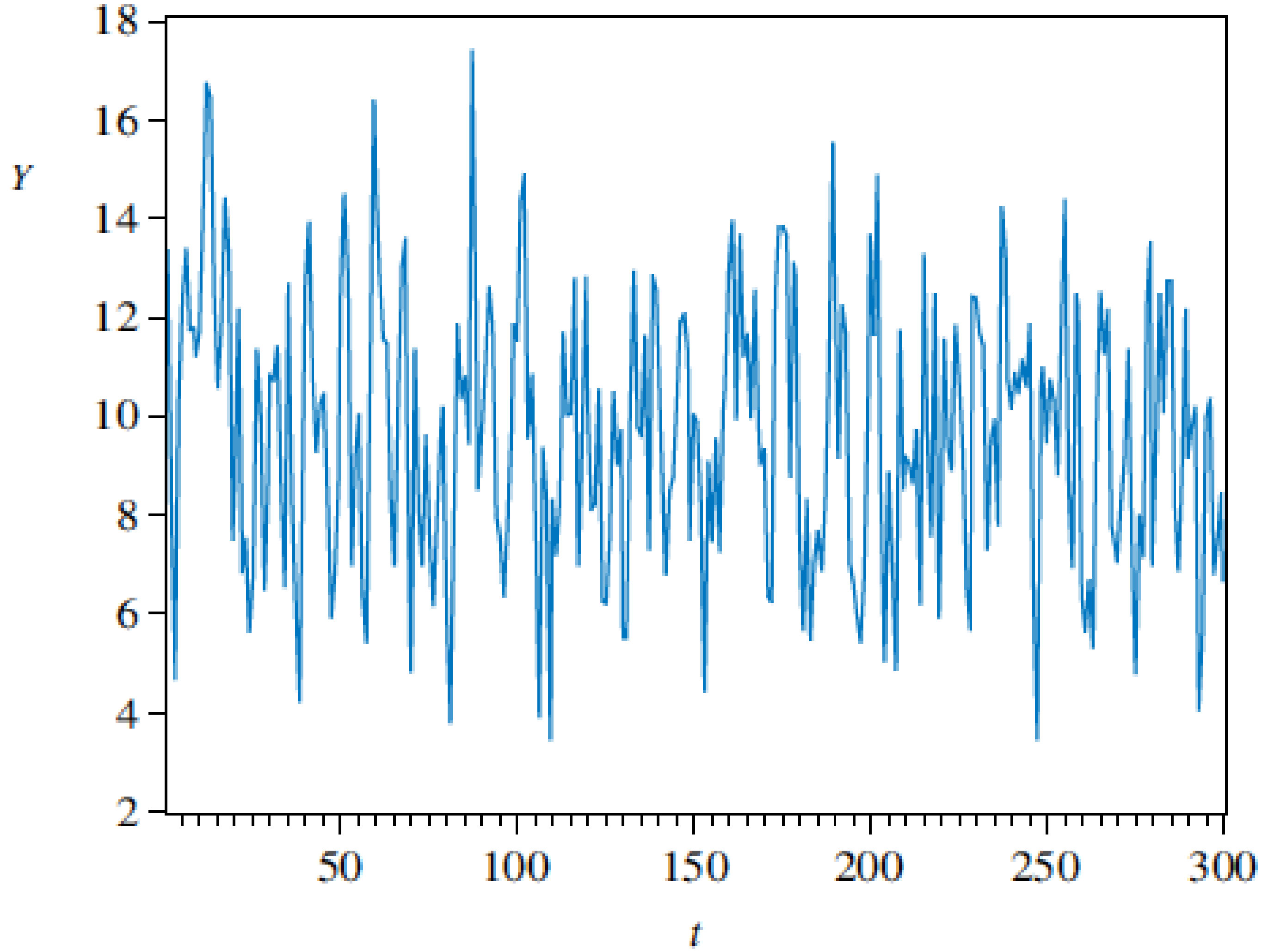


FIGURE 9.2 (b) time series of a nonstationary variable that is “slow-turning” or “wandering”

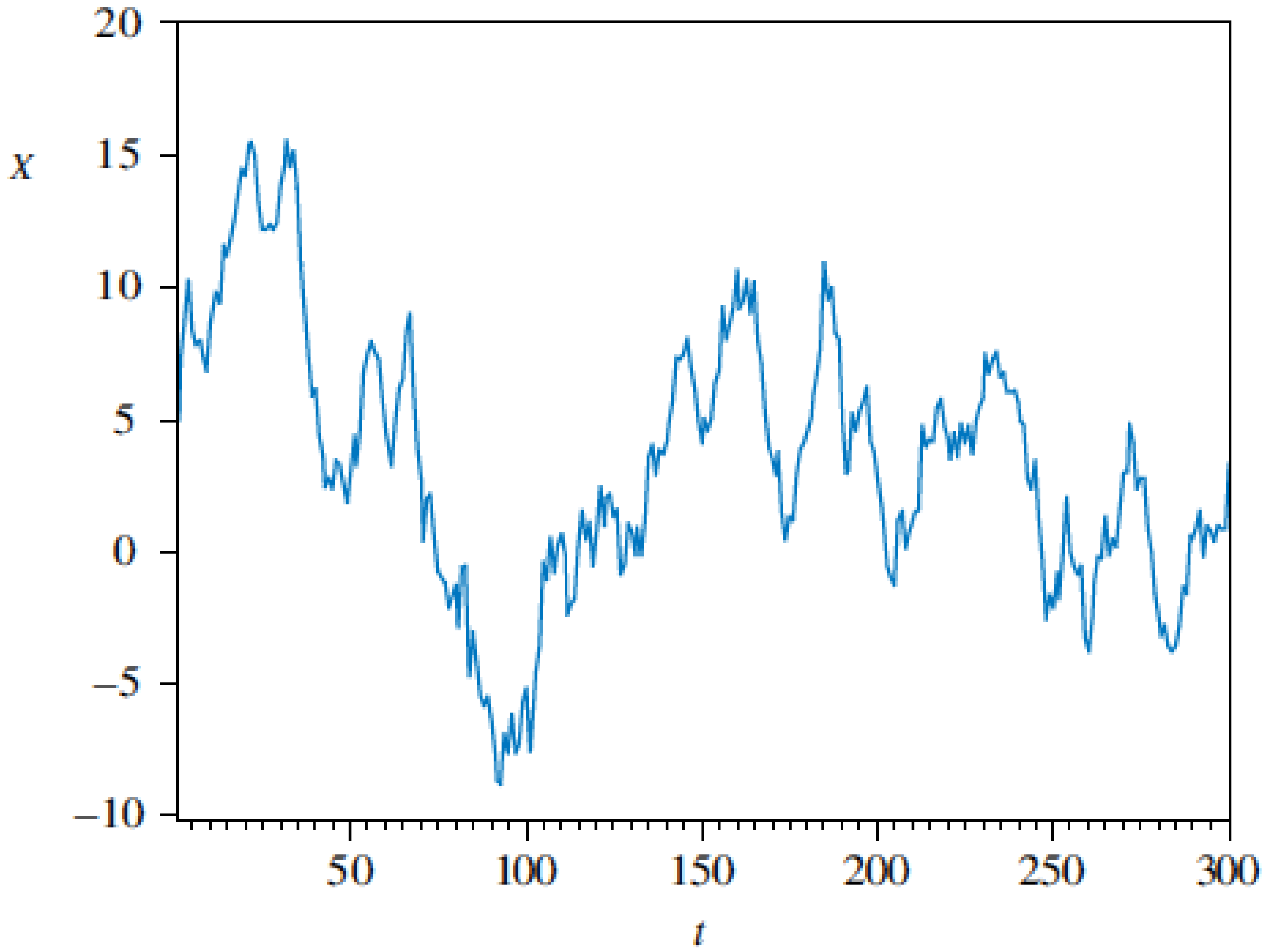
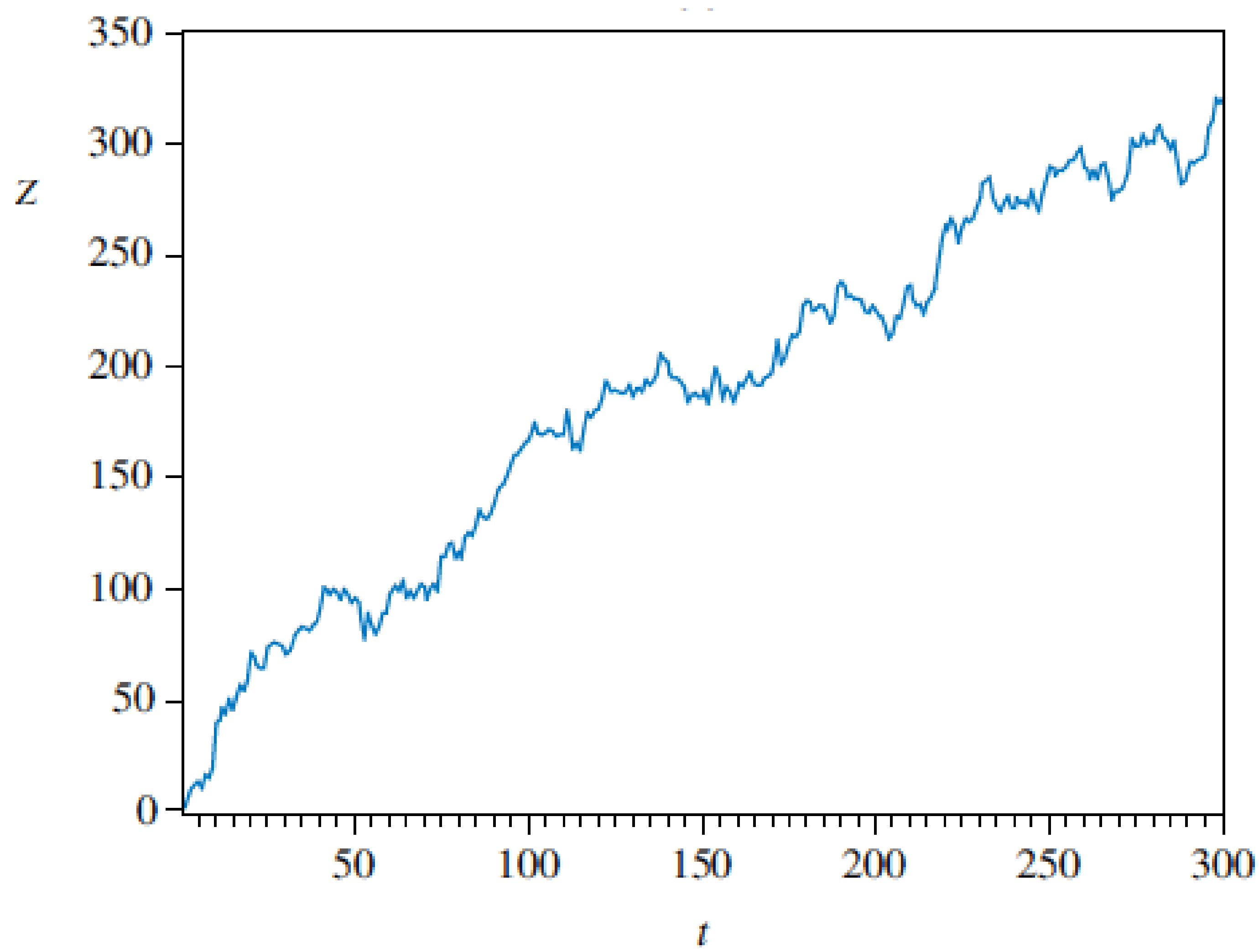


FIGURE 9.2 (c) time series of a nonstationary variable that "trends"



9.2 Serial Correlation

- When $\text{cov}(y_t, y_s) = 0$ for $t \neq s$ likely to be violated, and how do we assess its validity?
  - When a variable exhibits correlation over time, we say it is **autocorrelated** or **serially correlated**
    - These terms are used interchangeably

- Recall that the population correlation between two variables *x* and *y* is given by:

$$\rho_{xy} = \frac{\operatorname{cov}(x, y)}{\sqrt{\operatorname{var}(x)\operatorname{var}(y)}}$$

- For a stationary variable y_t we define

Eq. 9.5

$$\rho_1 = \frac{\text{cov}(y_t, y_{t-1})}{\sqrt{\text{var}(y_t)\,\text{var}(y_{t-1})}} = \frac{\text{cov}(y_t, y_{t-1})}{\text{var}(y_t)}$$

- The notation ρ_1 is used to denote the population correlation between observations that are one period apart in time
  - This is known also as the **population autocorrelation of order one**.
  - The second equality in Eq. 9.5 holds because $\text{var}(y_t) = \text{var}(y_{t-1})$, a property of time series that are stationary

- The first-order sample autocorrelation for *y* is obtained from Eq. 9.5 using the estimates:

$$\widehat{\text{cov}(y_t, y_{t-1})} = \frac{1}{T-1}\sum_{t=2}^{T}\left(y_t - \bar{y}\right)\left(y_{t-1} - \bar{y}\right)$$

$$\widehat{\text{var}(y_t)} = \frac{1}{T-1}\sum_{t=1}^{T}\left(y_t - \bar{y}\right)^2$$

9.2.1a
Computing
Autocorrelation

- Making the substitutions, we get:

Eq. 9.6

$$r_1 = \frac{\sum_{t=2}^{T}\left(y_t - \bar{y}\right)\left(y_{t-1} - \bar{y}\right)}{\sum_{t=1}^{T}\left(y_t - \bar{y}\right)^2}$$

9.2.1a
Computing Autocorrelation

- More generally, the ***k*-th order sample autocorrelation** for a series *y* that gives the correlation between observations that are *k* periods apart is:

Eq. 9.7

$$r_k = \frac{\sum_{t=k+1}^{T} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$$

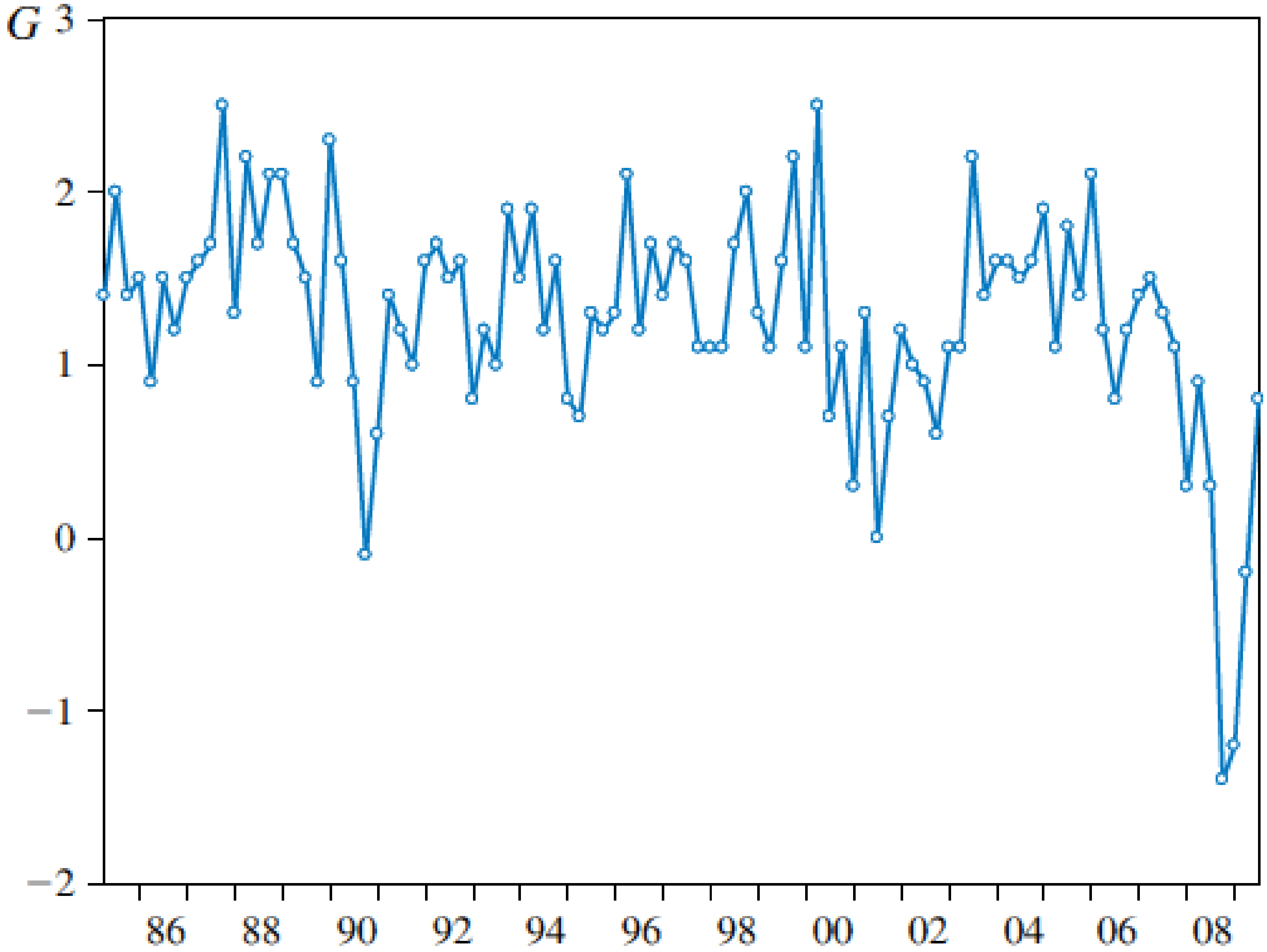
- How do we test whether an autocorrelation is significantly different from zero?
  - The null hypothesis is H_0: $\rho_k = 0$
  - A suitable test statistic is:

Eq. 9.8

$$Z = \frac{r_k - 0}{\sqrt{1/T}} = \sqrt{T}\, r_k \sim N(0,1)$$

FIGURE 9.3 Time series for U.S. GDP growth: 1985Q2 to 2009Q3

9.2.1b
An example: GDP growth rate



- Applying this to our problem, we get for the first four autocorrelations:

Eq. 9.9

$$r_1 = 0.494 \quad r_2 = 0.411 \quad r_3 = 0.154 \quad r_4 = 0.200$$

- For our problem, we have:

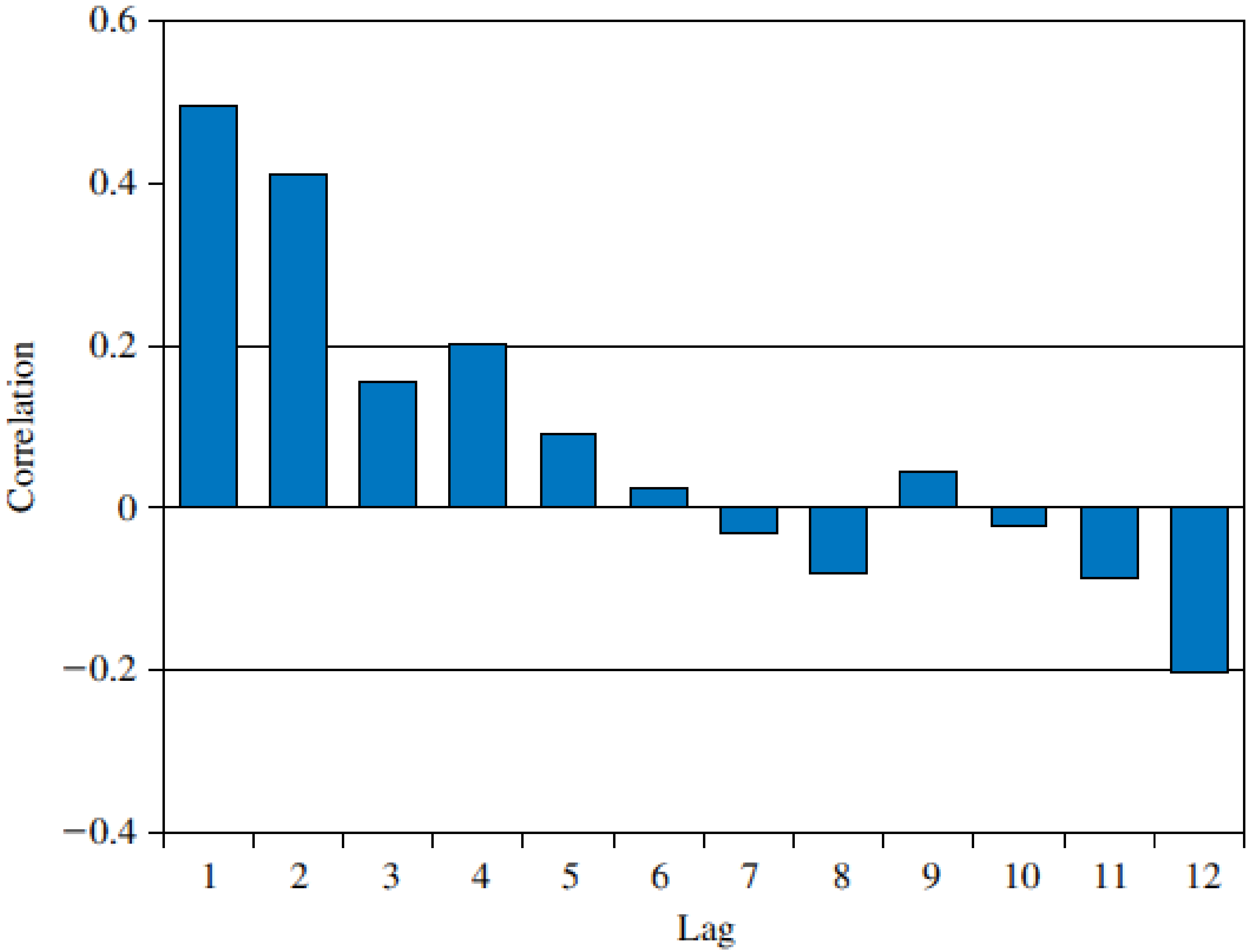
$$Z_1 = \sqrt{98} \times 0.494 = 4.89, \quad Z_2 = \sqrt{98} \times 0.414 = 4.10$$

$$Z_3 = \sqrt{98} \times 0.154 = 1.52, \quad Z_4 = \sqrt{98} \times 0.200 = 1.98$$

  – We reject the hypotheses $H_0: \rho_1 = 0$ and $H_0: \rho_2 = 0$
  – We have insufficient evidence to reject $H_0: \rho_3 = 0$
  – ρ_4 is on the borderline of being significant.
  – We conclude that the quarterly growth rate in U.S. GDP, exhibits significant serial correlation at lags one and two

- The **correlogram**, also called the **sample autocorrelation function**, is the sequence of autocorrelations $r_1, r_2, r_3, \ldots$
  - It shows the correlation between observations that are one period apart, two periods apart, three periods apart, and so on

FIGURE 9.4 Correlogram for GDP growth rate



- The correlogram can also be used to check whether the multiple regression assumption $\text{cov}(e_t, e_s) = 0$ for $t \neq s$ is violated

- Consider a model for a Phillips Curve:

Eq. 9.10

$$INF_t = INF_t^E - \gamma\left(U_t - U_{t-1}\right)$$

where INF_t is the inflation rate and U_t is the unemployment rate.

  – If we initially assume that inflationary expectations are constant over time ($\beta_1 = INF^E{}_t$) set $\beta_2 = -\gamma$, and add an error term:

Eq. 9.11

$$INF_t = \beta_1 + \beta_2 DU_t + e_t$$

FIGURE 9.5 (a) Time series for Australian price inflation

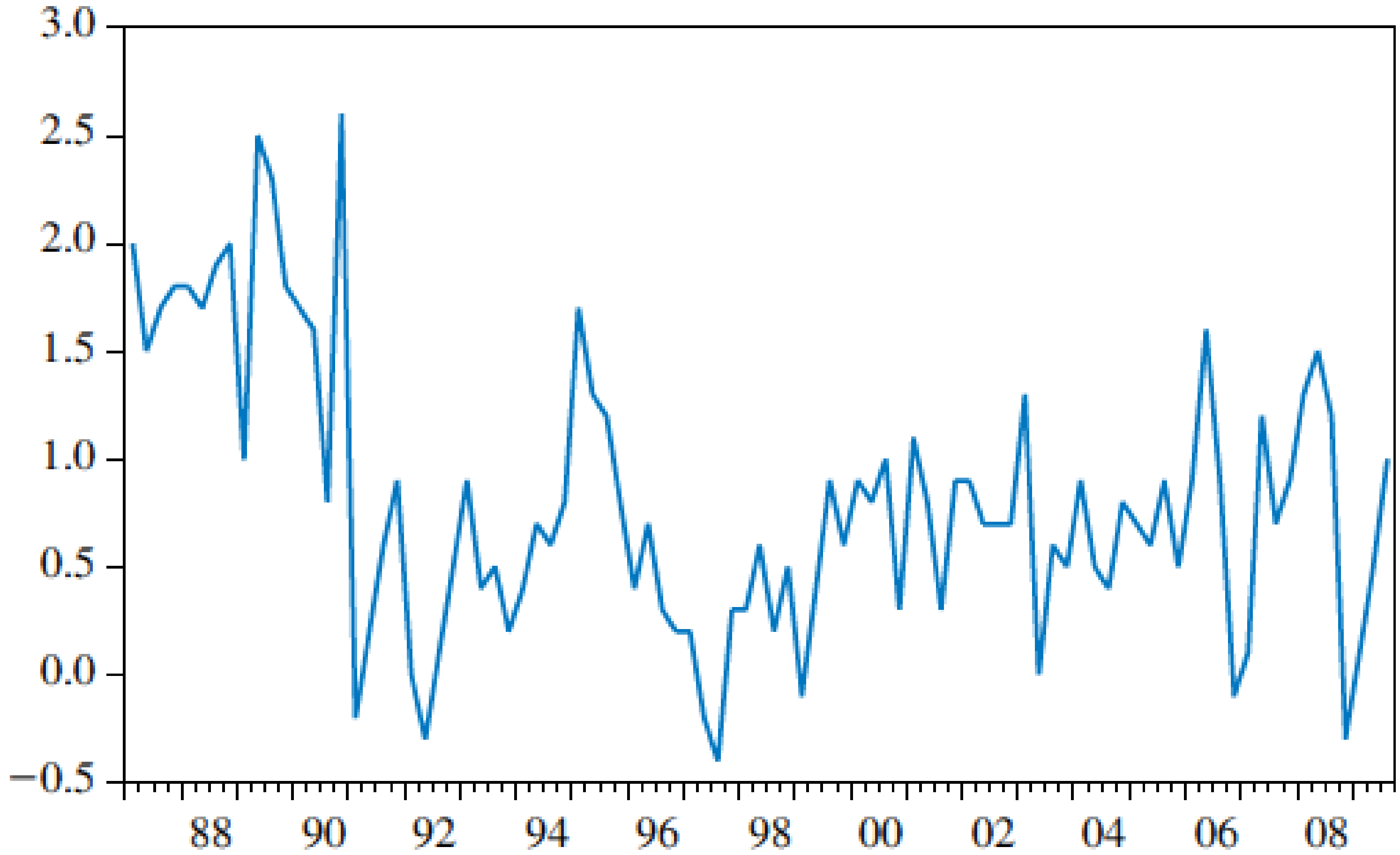
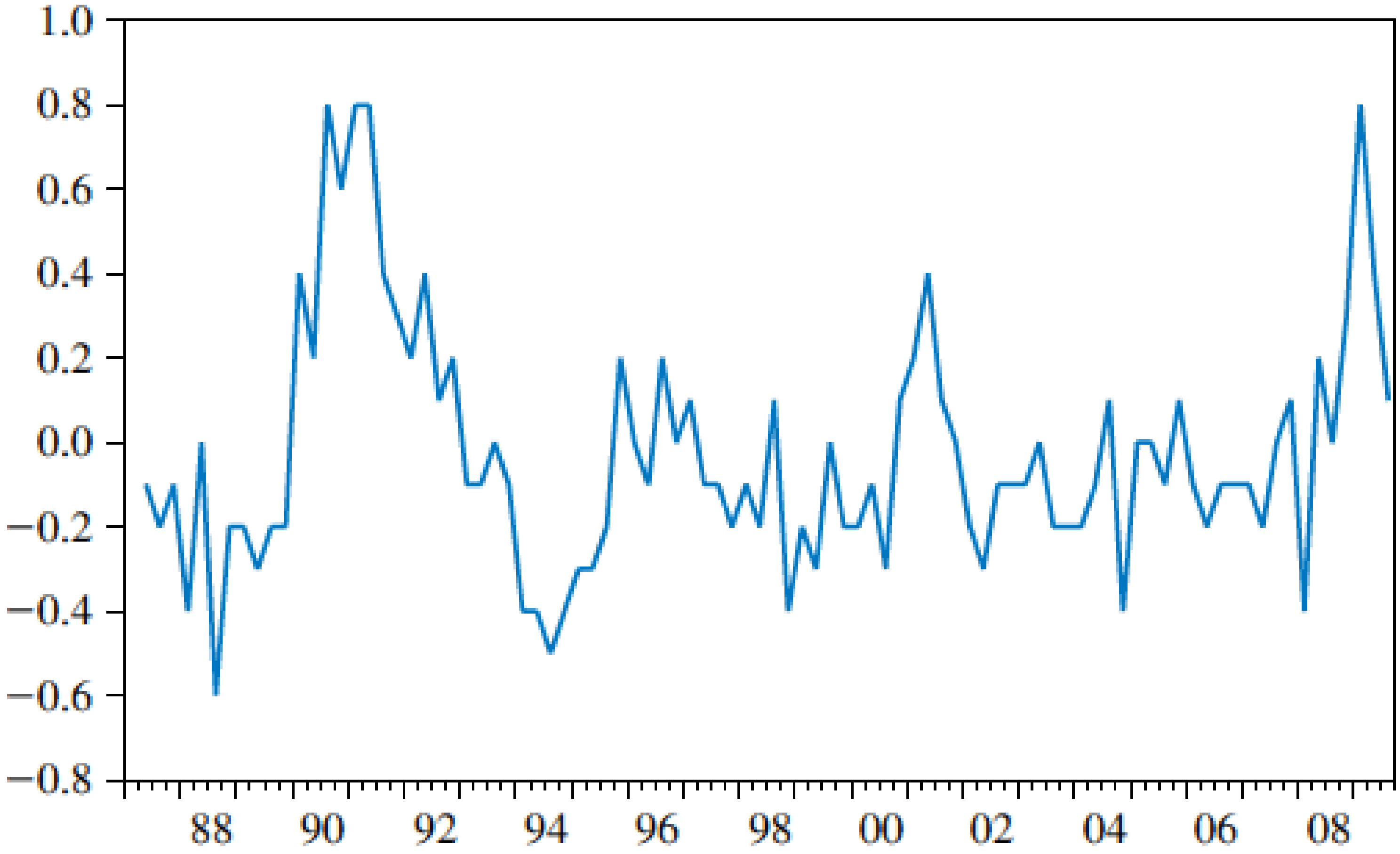


FIGURE 9.5 (b) Time series for the quarterly change in the Australian unemployment rate



- To determine if the errors are serially correlated, we compute the least squares residuals:

Eq. 9.20

$$\hat{e}_t = INF_t - b_1 - b_2 DU_t$$

- The k-th order autocorrelation for the residuals can be written as:

Eq. 9.12

$$r_k = \frac{\sum_{t=k+1}^{T} \hat{e}_t \hat{e}_{t-k}}{\sum_{t=1}^{T} \hat{e}_t^2}$$

– The least squares equation is:

Eq. 9.13

$$\widehat{INF} = 0.7776 - 0.5279 DU$$
$$(se) \quad (0.0658) \quad (0.2294)$$

- The values at the first five lags are:

$$r_1 = 0.549 \quad r_2 = 0.456 \quad r_3 = 0.433 \quad r_4 = 0.420 \quad r_5 = 0.339$$

- The test statistic is $Z_1 = \sqrt{90} \times 0.549 = 5.20$ which indicates that the null $\rho_1 = 0$ is rejected and the error term exhibits autocorrelation (of order 1).

9.3
Other Tests for Serially Correlated Errors

- An advantage of this test is that it readily generalizes to a joint test of correlations at more than one lag

- If e_t and e_{t-1} are correlated, then one way to model the relationship between them is to write:

Eq. 9.14

$$e_t = \rho e_{t-1} + v_t$$

  - We can substitute this into a simple regression equation:

Eq. 9.15

$$y_t = \beta_1 + \beta_2 x_t + \rho e_{t-1} + v_t$$

- We have one complication: e_t is unknown but we can use the least squares residuals $\hat{e}_t$ in place.
  - Also $\hat{e}_0$ is unknow. Two ways to handle this are:
    1. Delete the first observation and use a total of T-1 observations
    2. Set $\hat{e}_0 = 0$ and use all T observations

9.3.1
A Lagrange Multiplier Test

- To derive the relevant auxiliary regression for the autocorrelation LM test, we write the test equation as:

Eq. 9.16
$$y_t = \beta_1 + \beta_2 x_t + \rho \hat{e}_{t-1} + v_t$$

- But since we know that $y_t = b_1 + b_2 x_t + \hat{e}_t$, we get:

$$b_1 + b_2 x_t + \hat{e}_t = \beta_1 + \beta_2 x_t + \rho \hat{e}_{t-1} + v_t$$

Eq. 9.17

- Rearranging, we get:

$$\hat{e}_t = (\beta_1 - b_1) + (\beta_2 - b_2)x_t + \rho\hat{e}_{t-1} + v_t$$
$$= \gamma_1 + \gamma_2 x_t + \rho\hat{e}_{t-1} + v$$

  – If H_0: $\rho = 0$ is true, then LM = T x R^2 has an approximate $\chi^2_{(1)}$ distribution
    - T and R^2 are the sample size and goodness-of-fit statistic, respectively, from least squares estimation of Eq. 9.17

- Considering the two alternative ways to handle $\hat{e}_0$:

$$\text{(i)} \quad LM = (T-1) \times R^2 = 89 \times 0.3102 = 27.61$$

$$\text{(ii)} \quad LM = T \times R^2 = 90 \times 0.3066 = 27.59$$

  - These values are much larger than 3.84, which is the 5% critical value from a $\chi^2_{(1)}$-distribution
    - We reject the null hypothesis of no autocorrelation
  - Alternatively, we can reject H_0 by examining the *p*-value for LM = 27.61, which is 0.000

- For a four-period lag, we obtain:

$$\text{(iii)} \quad LM = (T-4) \times R^2 = 86 \times 0.3882 = 33.4$$

$$\text{(iv)} \quad LM = T \times R^2 = 90 \times 0.4075 = 36.7$$

– Because the 5% critical value from a $\chi^2_{(4)}$-distribution is 9.49, these LM values lead us to conclude that the errors are serially correlated

- This is used less frequently today because its critical values are not available in all software packages, and one has to examine upper and lower critical bounds instead
  - Also, unlike the LM and correlogram tests, its distribution no longer holds when the equation contains a lagged dependent variable
- When the error term is serially uncorrelated the D-W test statistic is equal to 2.
  - When the error term is positively serially correlated the D-W test statistic is between 0 and 2.
  - When the error term is negatively serially correlated the D-W test statistic is between 2 and 4.

9.4 Estimation with Serially Correlated Errors

- Three estimation procedures are considered:
  1. Least squares estimation
  2. An estimation procedure that is relevant when the errors are assumed to follow what is known as a first-order autoregressive model

     $$e_t = \rho e_{t-1} + v_t$$

  3. A general estimation strategy for estimating models with serially correlated errors

- Suppose we proceed with least squares estimation without recognizing the existence of serially correlated errors. What are the consequences?
  1. The least squares estimator is still a linear unbiased estimator, but it is no longer best
  2. The formulas for the standard errors usually computed for the least squares estimator are no longer correct
     - Confidence intervals and hypothesis tests that use these standard errors may be misleading

- It is possible to compute correct standard errors for the least squares estimator:
  - **HAC (heteroskedasticity and autocorrelation consistent) standard errors**, or **Newey-West standard errors**
    - These are analogous to the heteroskedasticity consistent standard errors

- Consider the model $y_t = \beta_1 + \beta_2 x_t + e_t$
  - The variance of b_2 is:

Eq. 9.18

$$\operatorname{var}(b_2) = \sum_t w_t^2 \operatorname{var}(e_t) + \sum\sum_{t \neq s} w_t w_s \operatorname{cov}(e_t, e_s)$$

$$= \sum_t w_t^2 \operatorname{var}(e_t) \left[1 + \frac{\sum\sum_{t \neq s} w_t w_s \operatorname{cov}(e_t, e_s)}{\sum_t w_t^2 \operatorname{var}(e_t)} \right]$$

where

$$w_t = (x_t - \bar{x}) \Big/ \sum_t (x_t - \bar{x})^2$$

- When the errors are not correlated, $\text{cov}(e_t, e_s) = 0$, and the term in square brackets is equal to one.
  - The resulting expression

$$\text{var}(b_2) = \sum_t w_t^2 \,\text{var}(e_t)$$

    is the one used to find heteroskedasticity-consistent (HC) standard errors
  - When the errors are correlated, the term in square brackets is estimated to obtain HAC standard errors

- Let's reconsider the Phillips Curve model:

Eq. 9.19

$$\widehat{INF} = 0.7776 - 0.5279DU$$
$$\begin{array}{lll} (0.0658) & (0.2294) & (\text{incorrect se}) \\ (0.1030) & (0.3127) & (\text{HAC se}) \end{array}$$

- The t and p-values for testing $H_0\!: \beta_2 = 0$ are:

$$t = -0.5279/0.2294 = -2.301 \qquad p = 0.0238 \qquad (\text{from LS standard errors})$$

$$t = -0.5279/0.3127 = -1.688 \qquad p = 0.0950 \qquad (\text{from HAC standard errors})$$

- Return to the Lagrange multiplier test for serially correlated errors where we used the equation:

Eq. 9.20

$$e_t = \rho e_{t-1} + v_t$$

  – Assume the v_t are uncorrelated random errors with zero mean and constant variances:

Eq. 9.21

$$E(v_t) = 0 \qquad \text{var}(v_t) = \sigma_v^2 \qquad \text{cov}(v_t, v_s) = 0 \quad \text{for } t \neq s$$

- Eq. 9.30 describes a **first-order autoregressive model** or a **first-order autoregressive process** for e_t
  - The term **AR(1) model** is used as an abbreviation for first-order autoregressive model
  - It is called an **autoregressive** model because it can be viewed as a regression model
  - It is called **first-order** because the right-hand-side variable is e_t lagged one period

- We assume that:

Eq. 9.22

$$-1 < \rho < 1$$

- The mean and variance of e_t are:

Eq. 9.23

$$E(e_t) = 0 \qquad \operatorname{var}(e_t) = \sigma_e^2 = \frac{\sigma_v^2}{1-\rho^2}$$

- The covariance term is:

Eq. 9.24

$$\operatorname{cov}(e_t, e_{t-k}) = \frac{\rho^k \sigma_v^2}{1-\rho^2}, \quad k > 0$$

- The correlation implied by the covariance is:

Eq. 9.25

$$\begin{aligned}\rho_k &= \operatorname{corr}(e_t, e_{t-k}) \\ &= \frac{\operatorname{cov}(e_t, e_{t-k})}{\sqrt{\operatorname{var}(e_t)\operatorname{var}(e_{t-k})}} \\ &= \frac{\operatorname{cov}(e_t, e_{t-k})}{\operatorname{var}(e_t)} \\ &= \frac{\rho^k \sigma_v^2 / (1-\rho^2)}{\sigma_v^2 / (1-\rho^2)} \\ &= \rho^k\end{aligned}$$

9.4.2a
Properties of an AR(1) Error

Eq. 9.26

- Setting $k = 1$:

$$\rho_1 = \text{corr}\left(e_t, e_{t-1}\right) = \rho$$

  - ρ represents the correlation between two errors that are one period apart
    - It is the **first-order autocorrelation** for *e*, sometimes simply called the autocorrelation coefficient
    - It is the population autocorrelation at lag one for a time series that can be described by an AR(1) model
    - r_1 is an estimate for ρ when we assume a series is AR(1)

Eq. 9.27

- Each e_t depends on all past values of the errors v_t:

$$e_t = v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \rho^3 v_{t-3} + \cdots$$

  - For the Phillips Curve, we find for the first five lags:

$$r_1 = 0.549 \quad r_2 = 0.456 \quad r_3 = 0.433 \quad r_4 = 0.420 \quad r_5 = 0.339$$

  - For an AR(1) model, we have:

$$\hat{\rho}_1 = \hat{\rho} = r_1 = 0.549$$

- For longer lags, we have:

$$\hat{\rho}_2 = \hat{\rho}^2 = (0.549)^2 = 0.301$$

$$\hat{\rho}_3 = \hat{\rho}^3 = (0.549)^3 = 0.165$$

$$\hat{\rho}_4 = \hat{\rho}^4 = (0.549)^4 = 0.091$$

$$\hat{\rho}_5 = \hat{\rho}^5 = (0.549)^5 = 0.050$$

- Our model with an AR(1) error is:

Eq. 9.28

$$y_t = \beta_1 + \beta_2 x_t + e_t \quad \text{with} \quad e_t = \rho e_{t-1} + v_t$$

with -1 < ρ < 1

– For the v_t, we have:

Eq. 9.29

$$E(v_t) = 0 \quad \text{var}(v_t) = \sigma_v^2 \quad \text{cov}(v_t, v_{t-1}) = 0 \quad \text{for } t \neq s$$

- With the appropriate substitutions, we get:

Eq. 9.30

$$y_t = \beta_1 + \beta_2 x_t + \rho e_{t-1} + v_t$$

  – For the previous period, the error is:

Eq. 9.31

$$e_{t-1} = y_{t-1} - \beta_1 - \beta_2 x_{t-1}$$

  – Multiplying by ρ:

Eq. 9.32

$$\rho e_{t-1} = e_t y_{t-1} - \rho\beta_1 - \rho\beta_2 x_{t-1}$$

- Substituting, we get:

Eq. 9.33

$$y_t = \beta_1(1-\rho) + \beta_2 x_t + \rho y_{t-1} - \rho\beta_2 x_{t-1} + v_t$$

- The coefficient of x_{t-1} equals $-\rho\beta_2$
  - Although Eq. 9.33 is a linear function of the variables x_t , y_{t-1} and x_{t-1}, it is not a linear function of the parameters (β_1, β_2, ρ)
  - The usual linear least squares formulas cannot be obtained by using calculus to find the values of (β_1, β_2, ρ) that minimize S_v
    - These are **nonlinear least squares estimates**

9.4.2b
Nonlinear Least Squares Estimation

- Our Phillips Curve model assuming AR(1) errors is:

Eq. 9.34

$$INF_t = \beta_1(1-\rho) + \beta_2 DU_t + \rho INF_{t-1} - \rho\beta_2 DU_{t-1} + v_t$$

  – Applying nonlinear least squares and presenting the estimates in terms of the original untransformed model, we have:

Eq. 9.35

$$\widehat{INF} = 0.7609 - 0.6944 DU \qquad e_t = 0.557 e_{t-1} + v_t$$
$$(se) \quad (0.1245) \quad (0.2479) \qquad\qquad (0.090)$$

- Nonlinear least squares estimation of Eq. 9.33 is equivalent to using an iterative generalized least squares estimator called the Cochrane-Orcutt procedure

- We have the model:

Eq. 9.36

$$y_t = \beta_1(1-\rho) + \beta_2 x_t + \rho y_{t-1} - \rho\beta_2 x_{t-1} + v_t$$

  – Suppose now that we consider the model:

Eq. 9.37

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

    • This new notation correspond to a general class of **autoregressive distributed lag (ARDL) models**

      – Eq. 9.37 is a member of this class

Eq. 9.38

- Note that Eq. 9.37 is the same as Eq. 9.36 since:

$$\delta = \beta_1(1-\rho) \quad \delta_0 = \beta_2 \quad \delta_1 = -\rho\beta_2 \quad \theta_1 = \rho$$

  – Eq. 9.36 is a restricted version of Eq. 9.37 with the restriction $\delta_1 = -\theta_1\delta_0$ imposed

9.4.3
Estimating a More General Model

- Applying the least squares estimator to Eq. 9.37 using the data for the Phillips curve example yields:

Eq. 9.39

$$\widehat{INF}_t = 0.3336 + 0.5593 INF_{t-1} - 0.6882 DU_t + 0.3200 DU_{t-1}$$
$$(se) \quad (0.0899) \quad (0.0908) \qquad (0.2575) \qquad (0.2499)$$

- The equivalent AR(1) estimates are:

$$\hat{\delta} = \hat{\beta}_1(1 - \hat{\rho}) = 0.7609 \times (1 - 0.5574) = 0.3368$$

$$\hat{\theta}_1 = \hat{\rho} = 0.5574$$

$$\hat{\delta}_0 = \hat{\beta}_2 = -0.6944$$

$$\hat{\delta}_1 = -\hat{\rho}\hat{\beta}_2 = -0.5574 \times (-0.6944) = 0.3871$$

- These are similar to our other estimates.
- One can test the validity of the AR(1) error term assumption by testing the hypothesis $\delta_1 = -\theta_1\delta_0$.

- When the simple regression model was introduced, we have assumed that the independent variable x is not random (SR5).
- We cannot maintain this assumption anymore in time-series models. Both y and x are sampled jointly so we do not know the value of x prior to sampling.
- We substitute SR5 with TSMR2 (see next page).
- This assumption in conjunction with the others guarantees that OLS estimators are unbiased and efficient.
- Adding TSMR6 the t and F-statistics follows the t and F distributions, respectively.

ASSUMPTIONS OF THE DISTRIBUTED LAG MODEL

TSMR1. $y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + e_t, \quad t = q+1, \ldots, T$

TSMR2. y and x are stationary random variables, and e_t is independent of current, past and future values of x.

TSMR3. $E(e_t) = 0$

TSMR4. $\text{var}(e_t) = \sigma^2$

TSMR5. $\text{cov}(e_t, e_s) = 0 \quad t \neq s$

TSMR6. $e_t \sim N(0, \sigma^2)$

- The ARDL model is more general than the previous specification. It includes a lagged dependent variable in the right-hand-side:

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

- The model can be written as:

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

- v_t is correlated with y_t so assumption TSMR2 is violated.
- This means that the least squares estimator is no longer unbiased but it **consistent** (under TSMR2A).
- The estimator is said to be consistent when it converges to the population parameter as the sample size tends to infinity.

ASSUMPTION FOR MODELS WITH A LAGGED DEPENDENT VARIABLE

TSMR2A In the multiple regression model $y_t = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_K x_{tK} + v_t$
Where some of the x_{tk} may be lagged values of y, v_t is uncorrelated with all x_{tk} and their past values.

Key Words

Keywords

- AR(1) error
- ARDL model
- autocorrelation
- Autoregressive distributed lags
- autoregressive error
- autoregressive model
- correlogram
- dynamic models
- finite distributed lag
- HAC standard errors
- lag operator
- lagged dependent variable
- LM test
- nonlinear least squares
- sample autocorrelations
- serial correlation
- T x R^2 form of LM test