Walter Enders, **Applied Econometric Time Series**. New York: John Wiley & Sons, Inc., 1995.

10. A MODEL OF THE WPI

The ARMA estimations performed in Section 8 were almost too straightforward. In practice, we rarely find a data series precisely conforming to a theoretical ACF or PACF. This section is intended to illustrate some of the ambiguities frequently encountered in the Box–Jenkins technique. These ambiguities may lead two equally skilled econometricians to estimate and forecast a series using very different ARMA processes. Many view the necessity to rely on the researcher's judgment and experience as a serious weakness of a procedure that is designed to be scientific.

It is useful to illustrate the Box–Jenkins modeling procedure by estimating a quarterly model of the U.S. Wholesale Price Index (WPI). The file labeled WPI.WK1 on the data disk contains the data used in this section. Exercise 10 at the end of this chapter will help you to reproduce the results reported below.

The top graph of Figure 2.5 clearly reveals that there is little point in modeling the series as being stationary; there is a decidedly positive trend or drift throughout the period 1960:I to 1990:IV. The first difference of the series seems to have a constant mean, although inspection of the middle graph suggests that the variance is an increasing function of time. As shown in the bottom graph of the same figure, the first difference of the logarithm (denoted by $\Delta lwpi$) is the most likely candidate to be covariance stationary. The large volatility of the WPI accompanying the oil price shocks in the 1970s should make us somewhat wary of the assumption that the process is covariance stationary. At this point, some researchers would make additional transformations intended to reduce the volatility exhibited in the 1970s. However, it seems reasonable to estimate a model of the $\{\Delta lwpi_t\}$ sequence. As always, you should maintain a healthy skepticism of the accuracy of your model.

Before reading on, you should examine the autocorrelation and partial autocorrelation functions of the $\{\Delta lwpi_t\}$ sequence shown in Figure 2.6. Try to identify the tentative models that you would want to estimate. In making your decision, note the following:

1. The ACF and PACF converge to zero reasonably quickly. We do not want to *overdifference* the data and try to model the $\{\Delta^2 lwpi_t\}$ sequence.
2. The theoretical ACF of a pure MA(q) process cuts off to zero at lag q and the theoretical ACF of an AR(1) model decays geometrically. Examination of the two graphs of Figure 2.6 suggests that neither of these specifications seems appropriate for the sample data.
3. The PACF is such that $\phi_{1,1} = 0.609$ and cuts off to 0.252 abruptly (i.e., $\phi_{2,2} = 0.252$). Overall, the PACF suggests that we should consider models such as $p = 1$ and $p = 2$. The ACF is suggestive of an AR(2) process or a process with both autoregressive and moving average components.
4. Note the jump in ACF at lag 4 and the small spike in the PACF at lag 4 ($\phi_{4,4} = 0.198$). Since we are using quarterly data, we might want to incorporate a seasonal factor at lag 4.

Figure 2.5 U.S. wholesale price index (1985 = 100).

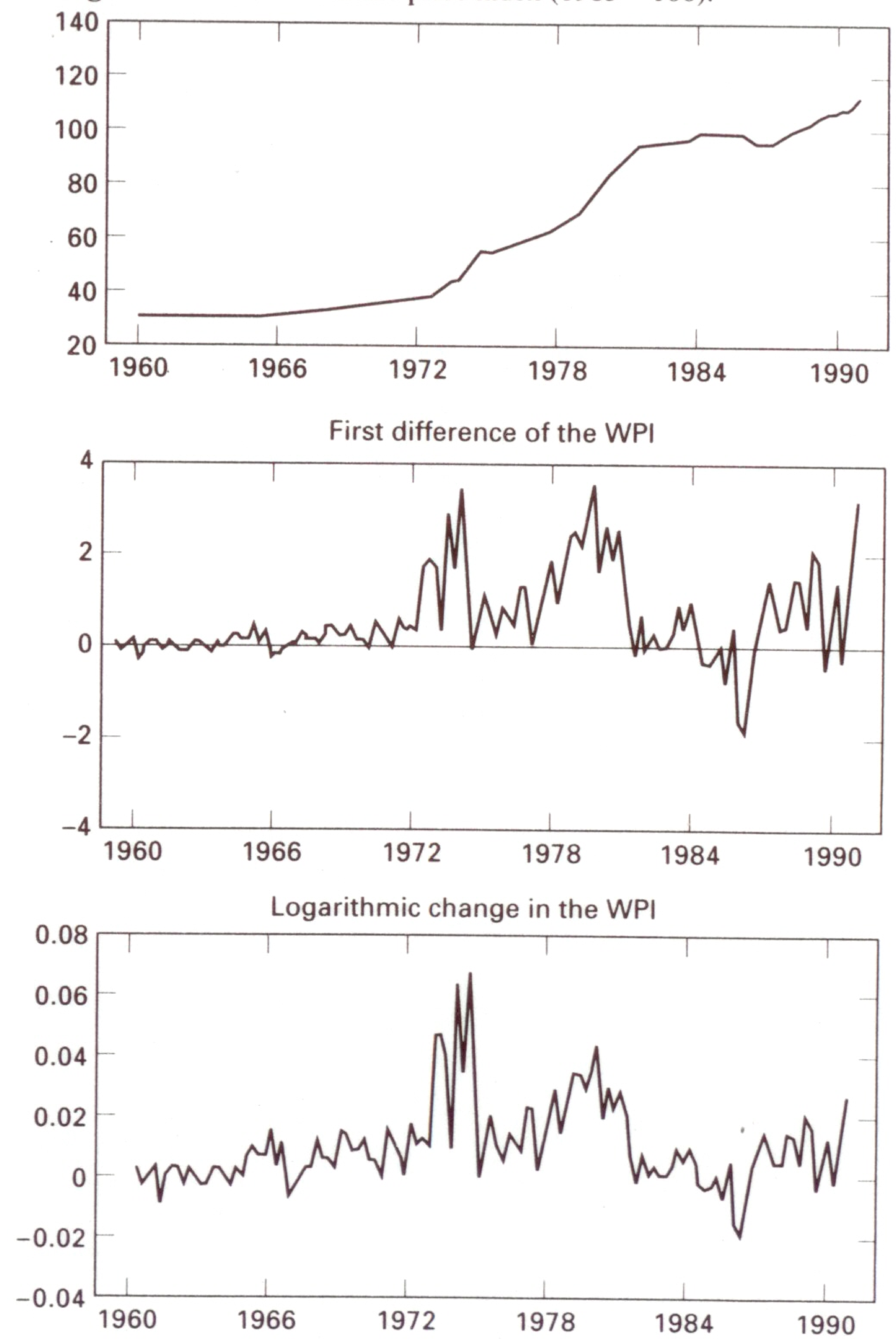
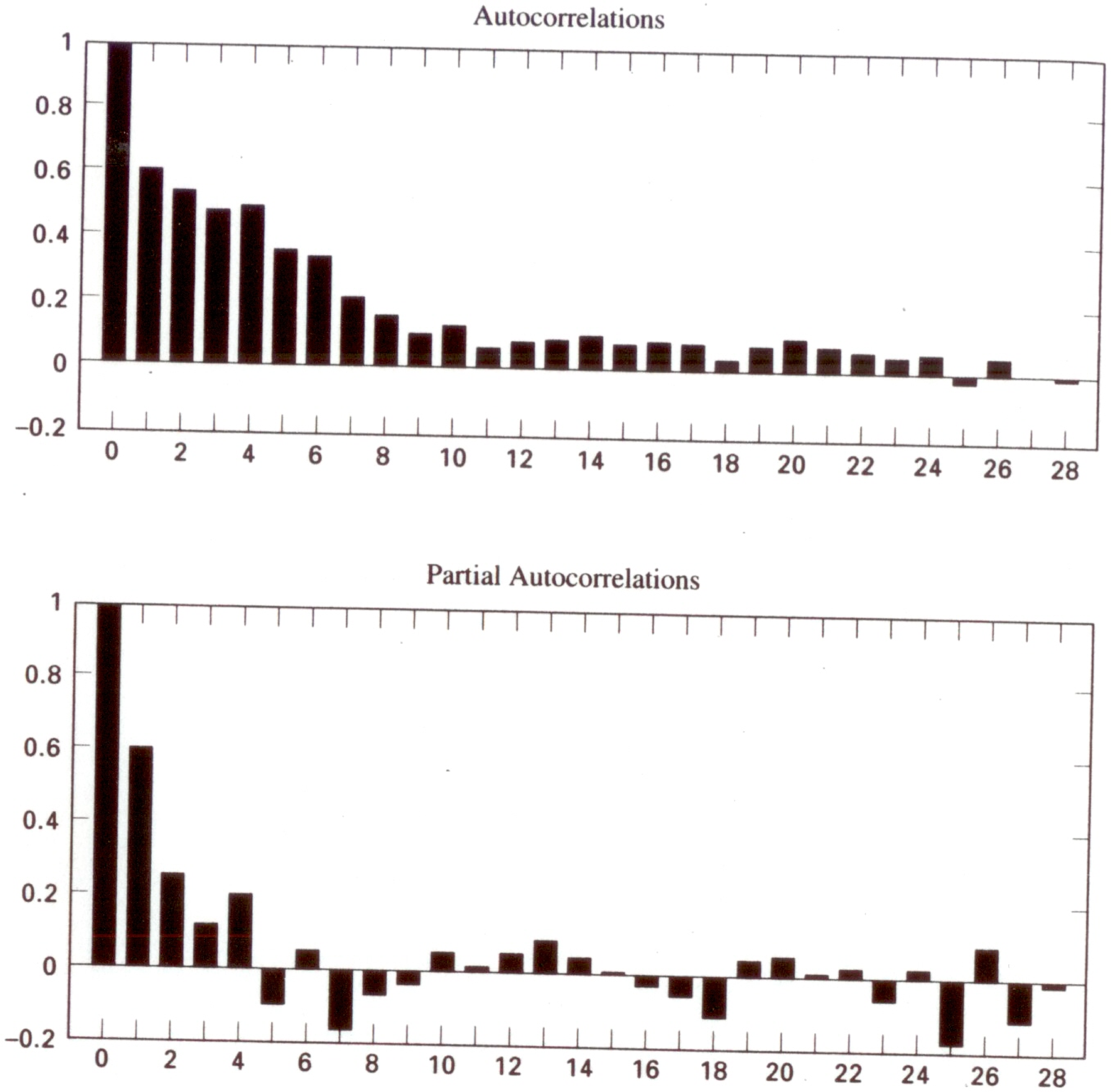


Figure 2.6 ACF and PACF for the logarithmic change in the WPI.



Points 1 to 4 suggest an ARMA(1, 1) or AR(2) model. In addition, we might want to consider models with a seasonal term at lag 4. Since computing time is inexpensive, we can estimate a variety of models and compare their results. Table 2.4 reports estimates of five tentative models; note the following points:

1. The estimated AR(1) model confirms our analysis in the identification stage. Although the estimated value of a_1 (0.618) is less than unity in absolute value and more than eight standard deviations from zero, the AR(1) specification is inadequate. Forming the Ljung–Box Q-statistic for 12 lags of the residuals yields a value of 23.6; we can reject the null that $Q = 0$ at the 1% significance level. Hence, the lagged residuals of this model exhibit substantial serial autocorrelation. Then we must eliminate this model from consideration.

2. The AR(2) model is an improvement over the AR(1) specification. The estimated coefficients ($a_1 = 0.456$ and $a_2 = 0.258$) are each significantly different from zero at the 1% level and imply characteristic roots in the unit circle. Q-statistics indicate that the autocorrelations of the residuals are not statistically significant. As measured by the AIC, the fit of the AR(2) model is superior to that of the AR(1); the SBC is the same for the two models. Overall, the AR(2) model dominates the AR(1) specification.
3. The ARMA(1, 1) specification dominates the AR(2) model. The estimated coefficients are of high quality (with t values of 14.9 and −4.22). The estimated value of a_1 is positive but less than unity, and the Q-statistics indicate that the autocorrelations of the residuals are not statistically significant. Moreover, all goodness-of-fit measures select the ARMA(1, 1) specification over the AR(2) model. Thus, there is little reason to maintain the AR(2) specification.

***Table 2.4:* Estimates of the WPI (Logarithmic First Differences)**

	$p = 1$ $q = 0$	$p = 2$ $q = 0$	$p = 1$ $q = 1$	$p = 1$ $q = 1, 4$	$p = 1$ $q = 2$
a_0	0.011 (4.14)	0.011 (3.31)	0.012 (2.63)	0.011 (2.76)	0.012 (2.62)
a_1	0.618 (8.54)	0.456 (5.11)	0.887 (14.9)	0.791 (9.21)	0.887 (13.2)
a_2		0.258 (2.89)			
β_1			−0.484 (−4.22)	−0.409 (−3.62)	−0.483 (−4.19)
β_2					−0.002 (−0.019)
β_4				0.315 (3.36)	
SSR	0.0156	0.0145	0.0141	0.0134	0.0141
AIC	−503.3	−506.1	−513.1	−518.2	−511.1
SBC	−497.7	−497.7	−504.7	−507.0	−499.9
$Q(12)$	23.6 (0.008)	11.7 (0.302)	11.7 (0.301)	4.8 (0.898)	11.7 (0.301)
$Q(24)$	28.6 (0.157)	15.6 (0.833)	15.4 (0.842)	9.3 (0.991)	15.3 (0.841)
$Q(30)$	40.1 (0.082)	22.8 (0.742)	22.7 (0.749)	14.8 (0.972)	22.6 (0.749)

Notes: Each coefficient is reported with the associated t-statistic for the null hypothesis that the estimated value is equal to zero.
SSR is the sum of squared residuals.
$Q(n)$ reports the Ljung–Box Q-statistic for the autocorrelations of the n residuals of the estimated model. With 122 observations, $T/4$ is approximately equal to 30. Significance levels are in parentheses.

4. In order to account for the possibility of seasonality, we estimated the ARMA(1, 1) model with an additional moving average coefficient at lag 4, that is, we estimated a model of the form $y_t = a_0 + a_1 y_{t-1} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_4 \epsilon_{t-4}$. More sophisticated seasonal patterns are considered in the next section. For now, note that the additive expression $\beta_4 \epsilon_{t-4}$ is often preferable to an additive autoregressive term of the form $a_4 y_{t-4}$. For truly seasonal shocks, the expression $\beta_4 \epsilon_{t-4}$ best captures spikes—not decay—at the quarterly lags. The coefficients of the estimated ARMA[1, (1, 4)] model are all highly significant with t-statistics of 9.21, –3.62, and 3.36.[8] The Q-statistics are all very low, implying that the autocorrelations of the residuals are statistically equal to zero. Moreover, the AIC and SBC strongly select this model over the ARMA(1, 1) model.
5. In contrast, the ARMA(1, 2) contains a superfluous coefficient. The t-statistic for β_2 is sufficiently low that we should eliminate this model.

Having identified and estimated a plausible model, we want to perform additional diagnostic checks of model adequacy. Due to the high volatility in the 1970s, the sample was split into the two subperiods: 1960:I to 1971:IV and 1972:I to 1990:IV. Model estimates for each subperiod are

$$\Delta lwpi_t = 0.004 + 0.641 \Delta lwpi_{t-1} + \epsilon_t - 0.351 \epsilon_{t-1} + 0.172 \epsilon_{t-4} \qquad (1960{:}\text{I–}1971{:}\text{IV})$$

and

$$\Delta lwpi_t = 0.016 + 0.753 \Delta lwpi_{t-1} + \epsilon_t - 0.394 \epsilon_{t-1} + 0.335 \epsilon_{t-4} \qquad (1972{:}\text{I–}1990{:}\text{IV})$$

The coefficients of the two models appear to be quite similar; we can formally test for the equality of coefficients using (2.47). Respectively, the sums of squared residuals for the two models are $SSR_1 = 0.001359$ and $SSR_2 = 0.011681$, and from Table 2.4 we can see that SSR = 0.0134. Since $T = 122$ and $n = 4$ (including the intercept means there are four estimated coefficients), (2.47) becomes

$$\begin{aligned} F &= [(0.0134 - 0.001359 - 0.011681)/4]/[0.001359 + 0.011681)/(122\text{-}8)] \\ &= 0.78681 \end{aligned}$$

With 4 degrees of freedom in the numerator and 114 in the denominator, we cannot reject the null of no structural change in the coefficients (i.e., we accept the hypothesis that there is no change in the structural coefficients).

As a final check, out-of-sample forecasts were constructed for each of the two models. By using additional data through 1992:II, the variance of the out-of-sample forecast errors of the ARMA(1, 1) and ARMA[1, (1,4)] models were calculated to be 0.00011 and 0.00008, respectively. Clearly, all the diagnostics select the ARMA[1, (1,4)] model. Although the ARMA[1, (1,4)] model appears to be adequate, other researchers might have selected a decidedly different model. Consider some of the alternatives listed below:

1. **Trends**: Although the logarithmic change of the WPI wholesale appears to be stationary, the ACF converges to zero rather slowly. Moreover, both the ARMA(1, 1) and ARMA[1, (1,4)] models yield estimated values of a_1 (0.887 and 0.791, respectively) that are close to unity. Some researchers might have chosen to model the second difference of the series. Others might have detrended the data using a deterministic time trend. Chapter 4 discusses formal tests for the appropriate form of the trend.
2. The seasonality of the data was modeled using a moving average term at lag 4. However, there are many other plausible ways to model the seasonality in the data, as discussed in the next section. For example, many computer programs are capable of estimating multiplicative seasonal coefficients. Consider the multiplicative seasonal model:

$$(1 - a_1 L) y_t = (1 + \beta_1 L)(1 + \beta_4 L^4) \epsilon_t$$

Here, the seasonal expression $\beta_4 \epsilon_{t-4}$ enters the model in a multiplicative, rather than a linear, fashion. Experimenting with various multiplicative seasonal coefficients might be a way to improve forecasting performance.

3. Given the volatility of the $\{\Delta lwpi_t\}$ sequence during the 1970s, the assumption of a constant variance might not be appropriate. Transforming the data using a square root, rather than the logarithm, might be more appropriate. A general class of transformations was proposed by Box and Cox (1964). Suppose that all values of $\{y_t\}$ are positive so that it is possible to construct the transformed $\{y_t^*\}$ sequence as

$$\begin{aligned} y_t^* &= (y_t^\lambda - 1)/\lambda, \qquad & \lambda \neq 0 \\ &= \ln(y_t), & \lambda = 0 \end{aligned}$$

The common practice is to transform the data using a preselected value of λ. Selecting a value of λ that is close to zero acts to "smooth" the sequence. As in the WPI example (which simply set $\lambda = 0$), an ARMA model can be fit to the transformed data. Although some software programs have the capacity to simultaneously estimate λ along with the other parameters of the ARMA model, this approach has fallen out of fashion. Instead, it is possible to actually model the variance using the methods discussed in Chapter 3.

11. SEASONALITY

Many economic processes exhibit some form of seasonality. The agricultural, construction, and travel sectors have obvious seasonal patterns resulting from their dependence on the weather. Similarly, the Thanksgiving–Christmas holiday season has a pronounced influence on the retail trade. In fact, the seasonal variation of some series may account for the preponderance of its total variance. Forecasts that ignore important seasonal patterns will have a high variance. In the last section, we saw how the inclusion of a four-quarter seasonal factor could help improve the model of the WPI. This section expands that discussion by illustrating some of the techniques that can be used to identify seasonal patterns.

Too many people fall into the trap of ignoring seasonality if they are working with **deseasonalized** or **seasonally adjusted** data. Suppose you collect a data set that the U.S. Bureau of the Census has "seasonally adjusted" using its X-11 method.[9] In principle, your seasonally adjusted data should have the seasonal pattern removed. However, caution is necessary. Although a standardized procedure may be necessary for a government agency reporting hundreds of series, the procedure might not be best for an individual wanting to model a single series. Even if you use seasonally adjusted data, a seasonal pattern might remain. This is particularly true if you do not use the entire span of data; the portion of the data used in your study can display more (or less) seasonality than the overall span. There is another important reason to be concerned about seasonality when using deseasonalized data. Implicit in any method of seasonal adjustment is a two-step procedure. First, the seasonality is removed, and second, the autoregressive and moving average coefficients are estimated using Box–Jenkins techniques. As surveyed in Bell and Hillmer (1984), often the seasonal and ARMA coefficients are best identified and estimated jointly. In such circumstances, it is wise to avoid using seasonally adjusted data.

Models of Seasonal Data

The Box–Jenkins technique for modeling seasonal data is no different from that of nonseasonal data. The twist introduced by seasonal data of period s is that the seasonal coefficients of the ACF and PACF appear at lags s, $2s$, $3s$, . . . , rather than at lags 1, 2, 3, ... For example, two purely seasonal models for quarterly data might be

$$y_t = a_4 y_{t-4} + \epsilon_t, \qquad |a_4| < 1 \tag{2.68}$$

and

$$y_t = \epsilon_t + \beta_4 \epsilon_{t-4} \tag{2.69}$$

You can easily convince yourself that the theoretical correlogram for (2.68) is such that $\rho_i = (a_4)^{i/4}$ if $i/4$ is an integer, and $\rho_i = 0$ otherwise; thus, the ACF exhibits decay at lags 4, 8, 12, ... For model (2.69), the ACF exhibits a single spike at lag 4 and all other correlations are zero.

In practice, identification will be complicated by the fact that the seasonal pattern will interact with the nonseasonal pattern in the data. The ACF and PACF for a combined seasonal/nonseasonal process will reflect both elements. Note that the final model of the wholesale price index estimated in the last section had the form

$$y_t = a_1 y_{t-1} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_4 \epsilon_{t-4} \tag{2.70}$$

Alternatively, an autoregressive coefficient at lag 4 might have been used to capture the seasonality:

$$y_t = a_1 y_{t-1} + a_4 y_{t-4} + \epsilon_t + \beta_1 \epsilon_{t-1} \tag{2.71}$$

Both these methods treat the seasonal coefficients additively; an AR or MA coefficient is added at the seasonal period. **Multiplicative seasonality** allows for the interaction of the ARMA and seasonal effects. Consider the multiplicative specifications:

$$(1 - a_1 L)y_t = (1 + \beta_1 L)(1 + \beta_4 L^4)\epsilon_t \tag{2.72}$$

$$(1 - a_1 L)(1 - a_4 L^4)y_t = (1 + \beta_1 L)\epsilon_t \tag{2.73}$$

Equation (2.72) differs from (2.70) in that it allows the moving average term at lag 1 to interact with the seasonal moving average effect at lag 4. In the same way, (2.73) allows the autoregressive term at lag 1 to interact with the seasonal autoregressive effect at lag 4. Many researchers prefer the multiplicative form since a rich interaction pattern can be captured with a small number of coefficients. Rewrite (2.72) as

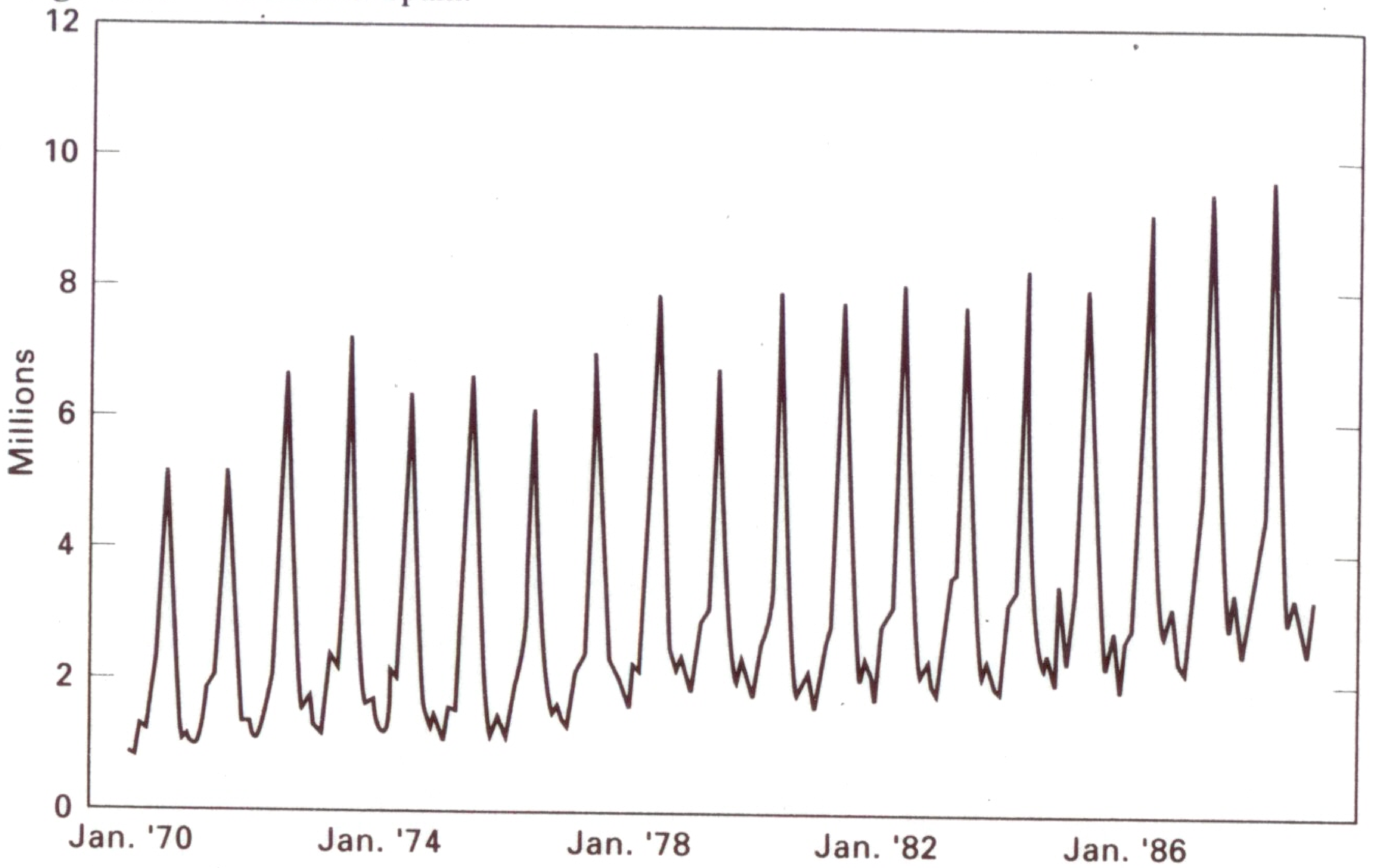
$$y_t = a_1 y_{t-1} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_4 \epsilon_{t-4} + \beta_1 \beta_4 \epsilon_{t-5} \tag{2.74}$$

Estimating only three coefficients (i.e., a_1, β_1, and β_4) allows us to capture the effects of an autoregressive term at lag 1 and the effects of moving average terms at lags 1, 4, and 5. Of course, you do not really get something for nothing. The estimates of the three moving average coefficients are interrelated. A researcher estimating the unconstrained model $y_t = a_1 y_{t-1} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_4 \epsilon_{t-4} + \beta_5 \epsilon_{t-5}$ would necessarily obtain a smaller residual sum of squares, since β_5 is not constrained to equal $\beta_1 \beta_4$. However, (2.72) is clearly the more parsimonious model. If the unconstrained value of β_5 approximates the product $\beta_1 \beta_4$, the multiplicative model will be preferable. For this reason, most software packages have routines capable of estimating multiplicative models. Otherwise, there are no theoretical grounds leading us to prefer one form of seasonality over another. As illustrated in the last section, experimentation and diagnostic checks are probably the best way to obtain the most appropriate model.

Seasonal Differencing

Spain is undoubtedly the most popular destination for European vacationers. During the months of July and August, the beaches along the Mediterranean coast swell with tourists basking in the sun. Figure 2.7 shows the monthly number of tourists visiting Spain between January 1970 and March 1989; the strong seasonal pattern dominates the movement in the series. You will also note that Spain's popularity has been growing; the series appears to be nonstationary in that the mean is increasing over time.

Figure 2.7 Tourism in Spain.



This combination of strong seasonality and nonstationarity is often found in economic data. The ACF for a nonstationary seasonal process is similar to that for a nonstationary nonseasonal process; with seasonal data the spikes at lags s, $2s$, $3s$, . . . do not exhibit rapid decay. The other autocorrelations are dwarfed by the seasonal effects. Notice ACF for the Spanish tourism data shown in Figure 2.8. The autocorrelation coefficients at lags 12, 24, 36, and 48 are all close to unity and the seasonal peaks decay slowly. The coefficients at lags 6, 18, 30, and 42 are all negative since tourism is always low 6 months from the summer boom.

Let y_t denote the log of number of tourists visiting Spain each month; the first step in the Box–Jenkins method is to difference the $\{y_t\}$ sequence so as to make it stationary. In contrast to the other series we examined, the appropriate way to difference strongly seasonal data is at the seasonal period. Formal tests for seasonal differencing are examined in Chapter 4. For now, it is sufficient to note that the seasonal difference $(1 - L^{12})y_t = y_t - y_{t-12}$ will have a smaller variance than the first difference $y_t - y_{t-1}$. In the Spanish data, the strong seasonality means that January-to-January and July-to-July changes are not as pronounced as the changes between June and July. Figure 2.9 shows the first and twelfth differences of the data; clearly, the twelfth difference has less variation and should be easier to identify and estimate.

The logarithmic twelfth difference (i.e., $y_t - y_{t-12}$) displays a flat ACF showing little tendency to decay. The first 12 of the autocorrelations are

ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ρ_7	ρ_8	ρ_9	ρ_{10}	ρ_{11}	ρ_{12}
0.26	0.31	0.26	0.28	0.23	0.24	0.19	0.21	0.19	0.20	0.15	–0.17

There is no reasonable way to fit a low-order model to the seasonally differenced data; the seasonal differencing did not eliminate the time-varying mean. In order to impart stationarity into the series, the next step is to take the first difference of the already seasonally differenced data. The ACF and PACF for the series $(1 - L)(1 - L^{12})y_t$ are shown in Figure 2.10; the properties of this series are much more amenable to the Box–Jenkins methodology. For the first 10 coefficients, the single spike in the ACF and uniform decay of the PACF suggest an MA(1) model. The significant coefficients at lags 11, 12, and 13 might result from additive or multiplicative seasonal factors. The estimates of the following three models are reported in Table 2.5:

$$(1 - L^{12})(1 - L)(1 - a_{12}L^{12})y_t = (1 + \beta_1 L)\epsilon_t \qquad \text{Model 1: Autoregressive}$$
$$(1 - L^{12})(1 - L)y_t = (1 + \beta_1 L)(1 + \beta_{12}L^{12})\epsilon_t \qquad \text{Model 2: Multiplicative moving average}$$
$$(1 - L^{12})(1 - L)y_t = (1 + \beta_1 L + \beta_{12}L^{12})\epsilon_t \qquad \text{Model 3: Additive moving average}$$

The point estimates of the coefficients all imply stationarity and invertibility. Moreover, all are at least six standard deviations from zero. However, the diagnostic statistics all suggest that model 2 is preferred. Model 2 has the best fit in that it has the lowest sum of squared residuals (SSR). Moreover, the Q-statistics for lags

Figure 2.8 Correlogram of tourism in Spain.

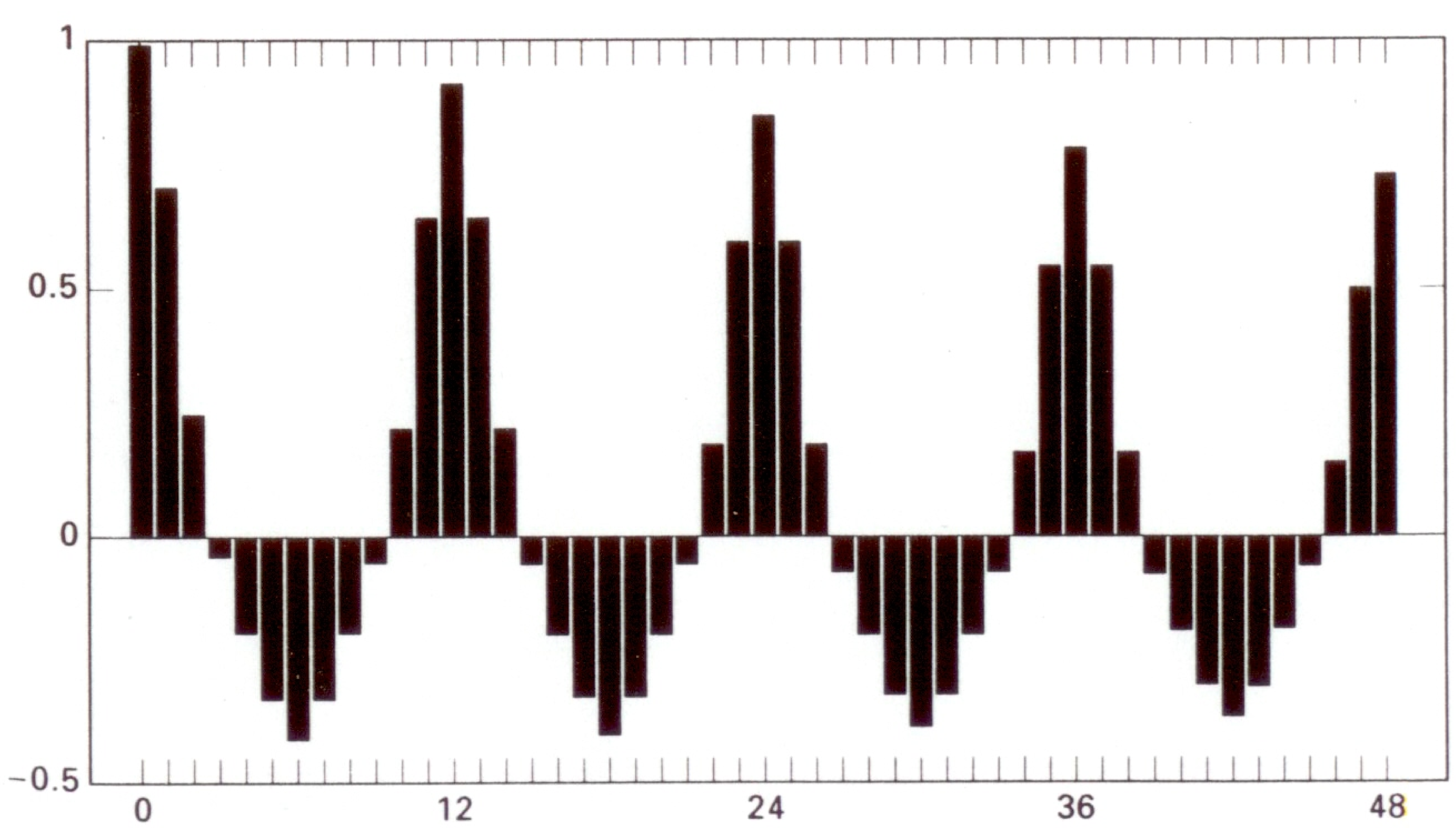


Figure 2.9 First and twelfth differences.

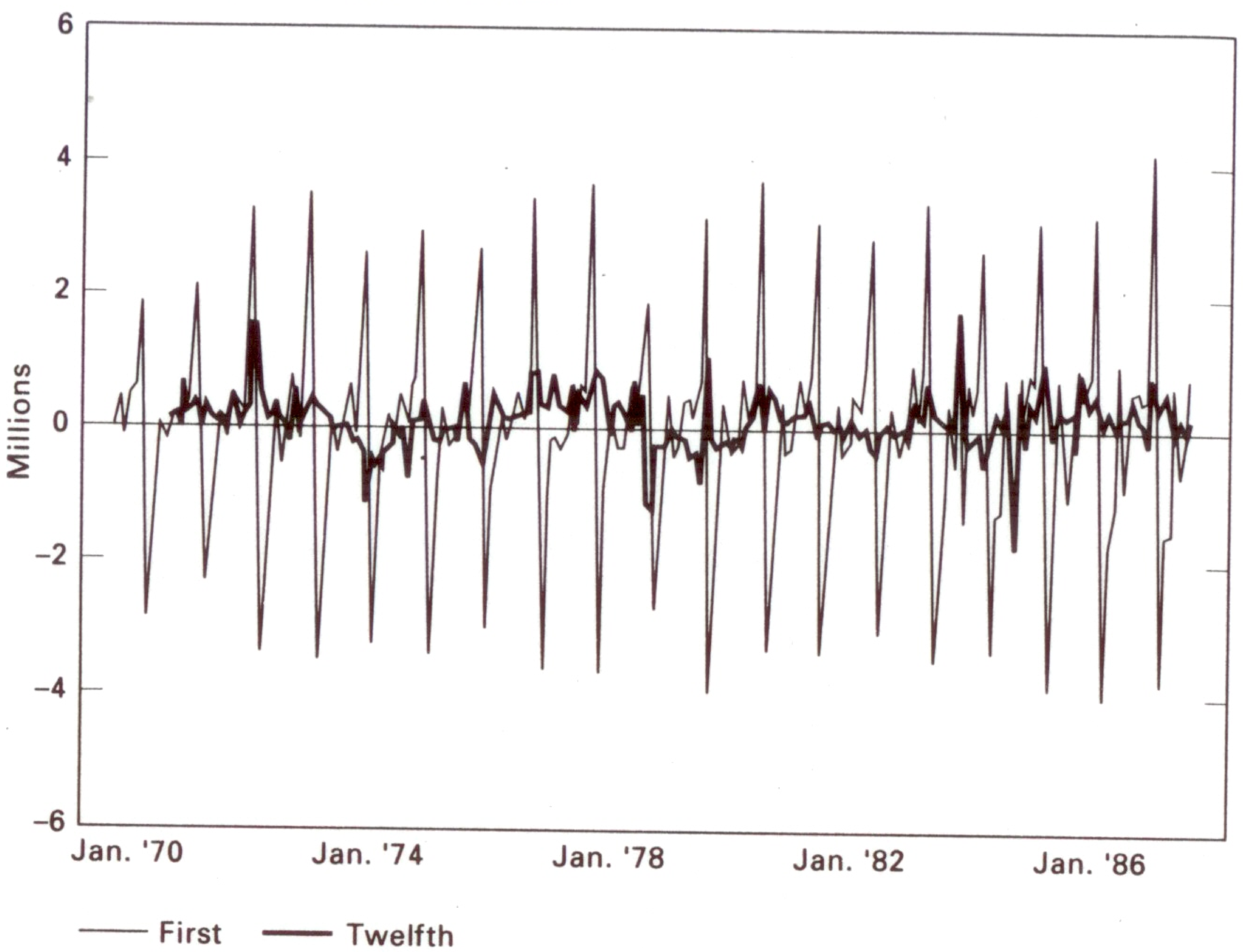


Table 2.5: **Three Models of Spanish Tourism**

	Model 1[1]	Model 2	Model 3
a_{12}	−0.408 (−6.54)		
β_1	−0.738 (−15.56)	0.740 (−16.14)	−0.640 (−14.75)
β_{12}		−0.671 (−13.02)	−0.306 (−7.00)
SSR	2.823	2.608	3.367
AIC	217.8	212.98	268.70
SBC	224.5	219.75	275.47
$Q(12)$	8.59 (0.571)	4.38 (0.928)	25.54 (0.004)
$Q(24)$	41.11 (0.007)	15.71 (0.830)	66.58 (0.000)
$Q(48)$	67.91 (0.019)	37.61 (0.806)	99.31 (0.000)

Clearly, there is no difference between an additive seasonality and multiplicative seasonality when all other autoregressive coefficients are zero.

12, 24, and 48 indicate that the residual autocorrelations are insignificant. In contrast, the residual correlations for model 1 are significant at long lags [i.e., $Q(24)$ and $Q(48)$ are significant at the 0.007 and 0.019 levels] and the residual correlations for model 3 are significant for lags 12, 24, and 48. Other diagnostic methods including overfitting and splitting the sample suggest that model 2 is appropriate.

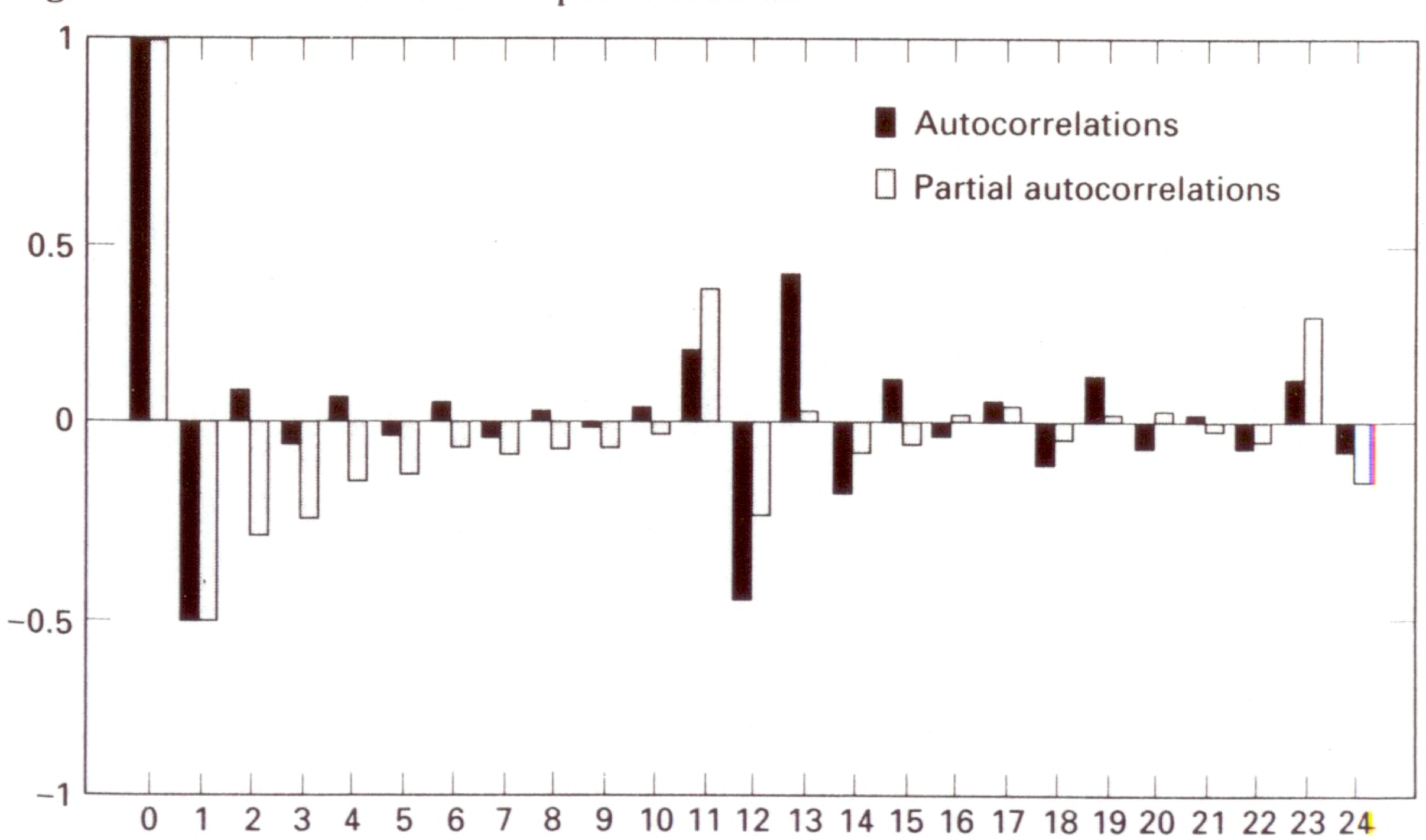
The procedures illustrated in this example of fitting a model to highly seasonal data are typical of many other series. With highly seasonal data, it is necessary to supplement the Box–Jenkins method:

1. In the identification stage, it is necessary to seasonally difference the data and check the ACF of the resultant series. Often, the seasonally differenced data will not be stationary. In such instances, the data may also need to be first-differenced.
2. Use the ACF and PACF to identify potential models. Try to estimate models with low-order nonseasonal ARMA coefficients. Consider both additive and multiplicative seasonality. Allow the appropriate form of seasonality to be determined by the various diagnostic statistics.

A compact notation has been developed that allows for the efficient representation of intricate models. As in previous sections, the dth difference of a series is denoted by Δ^d. For example,

$$\begin{aligned}\Delta^2 y_t &= \Delta(y_t - y_{t-1})\\ &= y_t - 2y_{t-1} + y_{t-2}\end{aligned}$$

Figure 2.10 ACF and PACF for Spanish Tourism.



Seasonally adjusted and first-differenced

A seasonal difference is denoted by Δ_s, where s is the period of the data. The Dth such seasonal difference is Δ_s^D. For example, if we wanted the second seasonal difference of the Spanish data, we could form

$$\begin{aligned}\Delta_{12}^2 y_t &= \Delta_{12}(y_t - y_{t-12}) \\ &= \Delta_{12} y_t - \Delta_{12} y_{t-12} \\ &= y_t - y_{t-12} - (y_{t-12} - y_{t-24}) \\ &= y_t - 2y_{t-12} + y_{t-24}\end{aligned}$$

Combining the two types of differencing yields $\Delta^d \Delta_s^D$. Multiplicative models are written in the form ARIMA$(p, d, q)(P, D, Q)_s$

where p and q = the nonseasonal ARMA coefficients
d = number of nonseasonal differences
P = number of multiplicative autoregressive coefficients;
D = number of seasonal differences
Q = number of multiplicative moving average coefficients
s = seasonal period

Using this notation, we can say that the fitted model of Spanish tourism is an ARIMA$(0, 1, 1)(0, 1, 1)_{12}$ model. In applied work, the ARIMA$(0, 1, 1)(0, 1, 1)_s$ model occurs routinely; it is called the "airline model" ever since Box and Jenkins (1976) used this model to analyze airline travel data.